

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:50:21 ; Search time 16939.5 Seconds
(without alignments)
11655.272 Million cell updates/sec

Title: US-10-017-410-1
Perfect score: 4175
Sequence: 1 ggacagagggtgtctcgat.....aaaaaaaaaaaaaaaaaac 4175

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4174	100.0	4174	10	AF282864 Mus muscu
2	3864.4	92.6	4078	10	BC059819 Mus muscu
3	3467.6	83.1	90370	10	BX005084 Mouse DNA
4	3467.6	83.1	207360	2	BX324222 Mus muscu
5	1222.4	29.3	184163	2	AC120245 Rattus no
6	1222.4	29.3	249734	2	AC097362 Rattus no
7	722.2	17.3	2852	6	C0842051 Sequence
8	722.2	17.3	2852	9	AK123581 Homo sapi
9	705	16.9	149462	9	AK017081 Homo sapi
10	653.8	15.7	822	6	AX206799 Sequence
11	510.8	12.2	531	11	G76816 S208P6606FG
12	400.4	9.6	595	6	AX400443 Sequence
13	382.6	9.2	784	11	BV053534 S212P6219
14	261.2	6.3	315	6	CQ736905 Sequence
15	261.2	6.3	315	6	CQ736906 Sequence
16	258	6.2	315	6	CQ744096 Sequence
17	232.4	5.6	2318	3	AY071232 Drosophil
18	229.2	5.5	1163	3	AF323976 Drosophil
19	205	4.9	1429	3	AK116177 Ciona int

20	183.6	4.4	163542	9	AL158206 Human DNA
21	183.6	4.4	219679	2	AC109451 Homo sapi
22	174.8	4.2	792	6	AX206797 Sequence
23	174.8	4.2	795	9	AF347024 Homo sapi
24	173.8	4.2	822	10	AF347023 Mus muscu
25	160.6	3.8	537	6	CQ605547 Sequence
26	131.2	3.1	1817	9	AK024898 Homo sapi
27	124	3.0	112713	9	AL391834 Human DNA
28	124	3.0	219679	2	AC109451 Homo sapi
29	114.2	2.7	358	6	AX419744 Sequence
30	114.2	2.7	358	6	AX980438 Sequence
31	114.2	2.7	358	6	BD115297 EST and e
32	108	2.6	2143	6	CQ492597 Sequence
33	108	2.6	2143	6	CQ492742 Sequence
34	108	2.6	2143	6	CQ493050 Sequence
35	92.6	2.2	4758	6	CQ605546 Sequence
36	92.6	2.2	13715	2	AC017332 Drosophil
37	92.6	2.2	78857	3	AC004364 Drosophil
38	92.6	2.2	157875	3	AE003665 Drosophil
39	92.6	2.2	164361	3	AC009252 Drosophil
40	92.6	2.2	165585	3	AC093044 Drosophil
41	78.4	1.9	257	6	CQ731034 Sequence
42	77.8	1.9	4071	6	CQ615131 Sequence
43	77.8	1.9	38138	2	AC017406 Drosophil
44	75.4	1.8	184065	2	AC108371 Pan trogl
45	72	1.7	173696	10	AL670603 Mouse DNA

ALIGNMENTS

RESULT 1
AF282864
LOCUS Mus musculus cancer related gene-liver 1 mRNA linear ROD 12-DEC-2001
DEFINITION AF282864 4174 bp mRNA
ACCESSION AF282864
VERSION AF282864.1 GI:17529683
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4174)
AUTHORS Graveel,C.R., Jatkoa,T., Madore,S.J., Holt,A.L. and Farnham,P.J.
TITLE Expression profiling and identification of novel genes in hepatocellular carcinomas
JOURNAL Oncogene 20 (21), 2704-2712 (2001)
MEDLINE 21313787
PubMed 11420682
REFERENCE 2 (bases 1 to 4174)
AUTHORS Graveel,C.R., Jatkoa,T., Madore,S.J., Holt,A.L. and Farnham,P.J.
TITLE Identification of genes deregulated in murine hepatocellular carcinomas using oligonucleotide microarrays and representational difference analysis
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 4174)
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2000) Oncology, University of Wisconsin, 1400 University Avenue, Madison, WI 53706, USA
FEATURES
source Location/Qualifiers
mol_type="mRNA"
db_xref="taxon:10090"
tissue_type="hepatocellular carcinoma"
35..862
/note="CRG-L1; putative"
/codon_start=1
/product="cancer related gene-liver 1"
/protein_id="AAL40408.1"
/db_xref="GI:17529684"
/translation="MGAPHWDLRAGSSEVDNEDNYTIVPAIEFYNTISNVLPFI"

LPPICMCLPROVATCFNSGYLIIWTLVVVGIGSVYFPHATLSFLQMLDELAILWVLM
CALAMWPRRVLPLPRNDRGRFKAIVCVLSAITTCGLAFIKPAINNIIMLIGLPCFA
LLVLEKRCNDVRVFKLGLFSGILSLALALFVWISDAQFCELLSSFFHPLHRCVWHILI
CLASYLGVCVCFAYFDAASEIPBQGPVIREFWPSEKMAFVGVPVSLCAHKHSPVKIT"

ORIGIN

Query Match		100.0%;	Score 4174;	DB 10;	Length 4174;		
Best Local Similarity		100.0%;	Pred No. 0;				
Matches 4174;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	GGCAGAGGCTGCTCCGATGCTCCAGAGCGGCATGGGCGCCCGCACTGTGGGACCA	60				
DB	1	GGCAGAGGCTGCTCCGATGCTCCAGAGCGGCATGGGCGCCCGCACTGTGGGACCA	60				
QY	61	CTGTGGGCTGCGAGTTCGGAGGTGATGGTGGAGGACCACTACACTATGTCGCTGC	120				
DB	61	CTGTGGGCTGCGAGTTCGGAGGTGATGGTGGAGGACCACTACACTATGTCGCTGC	120				
QY	121	CATTGCCGAGTTCTACACACGATCAGCAACGTCCTGTGTTTTCATTTTACCTCCCATCG	180				
DB	121	CATTGCCGAGTTCTACACACGATCAGCAACGTCCTGTGTTTTCATTTTACCTCCCATCG	180				
QY	181	CATGTGCTGTTCCGCCAGTAGCGAACGTCCTCAACAGCGCATCTACTTAATATGGAC	240				
DB	181	CATGTGCTGTTCCGCCAGTAGCGAACGTCCTCAACAGCGCATCTACTTAATATGGAC	240				
QY	241	GCTCCTAGTTGAGTGGGATTTGGATCTGTCTACTTCCATGCCAAGCTGAGTTTCTGGG	300				
DB	241	GCTCCTAGTTGAGTGGGATTTGGATCTGTCTACTTCCATGCCAAGCTGAGTTTCTGGG	300				
QY	301	TCAGATGCTGATGAATCTGCCATTTCTGTGGGTTCTGATGTGCTTTGGCCATGTGTT	360				
DB	301	TCAGATGCTGATGAATCTGCCATTTCTGTGGGTTCTGATGTGCTTTGGCCATGTGTT	360				
QY	361	TCCAGAGGATTTTACCAAGATCTTTCGGAATGACAGGGGACAGTTTCAAGGCAGTGT	420				
DB	361	TCCAGAGGATTTTACCAAGATCTTTCGGAATGACAGGGGACAGTTTCAAGGCAGTGT	420				
QY	421	GTGTGCTGCTGCTCAATTTACAACTGCTTGGCGTTTATCAGCCCGGCATCAATAAT	480				
DB	421	GTGTGCTGCTGCTCAATTTACAACTGCTTGGCGTTTATCAGCCCGGCATCAATAAT	480				
QY	481	TTCCCTGATGATCTGGGACTTCCAATGCACTGCGCTGCTTGTGAGAGCTGAAGAGTG	540				
DB	481	TTCCCTGATGATCTGGGACTTCCAATGCACTGCGCTGCTTGTGAGAGCTGAAGAGTG	540				
QY	541	TGACAAATGCTGCTGTTTAACTGGGCTCTTCTCTGGGCTCTGGTGGACTCTGGCTCT	600				
DB	541	TGACAAATGCTGCTGTTTAACTGGGCTCTTCTCTGGGCTCTGGTGGACTCTGGCTCT	600				
QY	601	CTTCTGCTGATCAGCAAGCCTTCTGTGAGCTGCTCTCTCTGAGCTGCTCTGCTCT	660				
DB	601	CTTCTGCTGATCAGCAAGCCTTCTGTGAGCTGCTCTCTCTGAGCTGCTCTGCTCT	660				
QY	661	CCTGCACTGTGTGGCATATTTCTCATCTGCTTCTGTAACCTGGGCTGTGTGCTT	720				
DB	661	CCTGCACTGTGTGTGGCATATTTCTCATCTGCTTCTGTAACCTGGGCTGTGTGCTT	720				
QY	721	GGCTACTTTGATGCTGCTCAGAGATACCTGAGCAAGGTCAGTCAATCTGGCC	780				
DB	721	GGCTACTTTGATGCTGCTCAGAGATACCTGAGCAAGGTCAGTCAATCTGGCC	780				
QY	781	CAGCGAGAAATGGGCTTTTATTTGTGTGCTTATGTGTCCTTCTGTGTCGCCCAAGAA	840				
DB	781	CAGCGAGAAATGGGCTTTTATTTGTGTGCTTATGTGTCCTTCTGTGTCGCCCAAGAA	840				
QY	841	GTGCGCATCAAGATCACTGATGGCAAGGAGTGAACAGCTTCTACTTACTTCTATT	900				
DB	841	GTGCGCATCAAGATCACTGATGGCAAGGAGTGAACAGCTTCTACTTACTTCTATT	900				
QY	901	CGAGTGGCGCTGGCTTCTGTTGCTAGCAAGATGCTGAGGGGTTGAGGAATTTGGTG	960				
DB	901	CGAGTGGCGCTGGCTTCTGTTGCTAGCAAGATGCTGAGGGGTTGAGGAATTTGGTG	960				

QY	961	TGCTGTGGGTGTTTAAAAATTCCTCTCTTTGTGATCTAACTGGACACTGTGCTGACCT	1020				
DB	961	TGCTGTGGGTGTTTAAAAATTCCTCTCTTTGTGATCTAACTGGACACTGTGCTGACCT	1020				
QY	1021	CCCTAGGTTAAGTACAGGGCTCAGGACATAGGTGTCTTCTCAGTATCTTTCCAGGC	1080				
DB	1021	CCCTAGGTTAAGTACAGGGCTCAGGACATAGGTGTCTTCTCAGTATCTTTCCAGGC	1080				
QY	1081	ATACGGGCTTGTCTGGGTATGTCTCATATCAATCAACAGAGTAGTCTTTGGGAGACCT	1140				
DB	1081	ATACGGGCTTGTCTGGGTATGTCTCATATCAATCAACAGAGTAGTCTTTGGGAGACCT	1140				
QY	1141	AGGCAACCCAAAGTTTCTGTGAGAGGGTAGTGTCTTCCACCAATTTCCAGACT	1200				
DB	1141	AGGCAACCCAAAGTTTCTGTGAGAGGGTAGTGTCTTCCACCAATTTCCAGACT	1200				
QY	1201	CTCAACCCCTCATAAACAGACCTTCTGATTTGGTGATGAAGGTTCCAGAACTTTTCA	1260				
DB	1201	CTCAACCCCTCATAAACAGACCTTCTGATTTGGTGATGAAGGTTCCAGAACTTTTCA	1260				
QY	1261	TTTTGCCGGGAGAACTGTCTCTTCAAAAAACAAAGTGGGCAAAAACACGTGTGGGGGTG	1320				
DB	1261	TTTTGCCGGGAGAACTGTCTCTTCAAAAAACAAAGTGGGCAAAAACACGTGTGGGGGTG	1320				
QY	1321	TGACTAAGACGGGTGGCTTGTCTGCACTCTTGTAGCTTTCCAGGTTTCTCCAGTGT	1380				
DB	1321	TGACTAAGACGGGTGGCTTGTCTGCACTCTTGTAGCTTTCCAGGTTTCTCCAGTGT	1380				
QY	1381	TGTGATCTTCCACTACTCTGCTGAGAGTGGAGCCACGGCTTCAGAGGCAAGCTG	1440				
DB	1381	TGTGATCTTCCACTACTCTGCTGAGAGTGGAGCCACGGCTTCAGAGGCAAGCTG	1440				
QY	1441	GCAACACCTCTATGCCAAAAGCTACACTCTCTTTAAAGCACACATTTACATAGACATAT	1500				
DB	1441	GCAACACCTCTATGCCAAAAGCTACACTCTCTTTAAAGCACACATTTACATAGACATAT	1500				
QY	1501	TTCTGCTCTTCCAGAGTGCAGAGCTCAGACCCACAGAGAACTCTTCCAGGTTATGTGAA	1560				
DB	1501	TTCTGCTCTTCCAGAGTGCAGAGCTCAGACCCACAGAGAACTCTTCCAGGTTATGTGAA	1560				
QY	1561	GATTCACACACAGCCCTCTGTGACCTCTGTGAGAGCAAGGCTCGTTGTAATCGG	1620				
DB	1561	GATTCACACACAGCCCTCTGTGACCTCTGTGAGAGCAAGGCTCGTTGTAATCGG	1620				
QY	1621	GAGGAGCTCTGGAGGCTTGACTGTGAGCGCTGTGAAGATTTTGGAAAAGCTTTCTTA	1680				
DB	1621	GAGGAGCTCTGGAGGCTTGACTGTGAGCGCTGTGAAGATTTTGGAAAAGCTTTCTTA	1680				
QY	1681	TTGGAAGCTTTTGTGTGCTTACTCTCAAGATTTCTGACCCCGTTTCTGCTGTTTTT	1740				
DB	1681	TTGGAAGCTTTTGTGTGCTTACTCTCAAGATTTCTGACCCCGTTTCTGCTGTTTTT	1740				
QY	1741	AGGGATATGTGCTTCTGAGTGCAGTGCCTCTGAGTGTGGGCTTCGGGCTTTTGT	1800				
DB	1741	AGGGATATGTGCTTCTGAGTGCAGTGCCTCTGAGTGTGGGCTTCGGGCTTTTGT	1800				
QY	1801	TTTCATAATACCTGGTGTCTTCTCAGAGCACTTTGCTCTTCTGATGCTGGGC	1860				
DB	1801	TTTCATAATACCTGGTGTCTTCTCAGAGCACTTTGCTCTTCTGATGCTGGGC	1860				
QY	1861	GGTCACTACATCTGATTTGGCTGGGATGCCAGGAGAAAGGGGACAGCTCTGACA	1920				
DB	1861	GGTCACTACATCTGATTTGGCTGGGATGCCAGGAGAAAGGGGACAGCTCTGACA	1920				
QY	1921	TGAATGTCTCTCAGCTGCATCTACTGTCTTCTCAGTGGCTCTCTTCAAGGTATCAAGTGG	1980				
DB	1921	TGAATGTCTCTCAGCTGCATCTACTGTCTTCTCAGTGGCTCTCTTCAAGGTATCAAGTGG	1980				
QY	1981	GTATGGGATATCTCAGGATGTTTGTAACTCAGGACCTTCTGCTTTCTGACATTTCCAAT	2040				
DB	1981	GTATGGGATATCTCAGGATGTTTGTAACTCAGGACCTTCTGCTTTCTGACATTTCCAAT	2040				

QY 2041 GTCACTGGTGAGACTGCACCTCAGGGGACTGGAGTTGGAAACCGTTCTGAGTGGTGCTCT 2100
DB |||||
QY 2041 GTCACTGGTGAGACTGCACCTCAGGGGACTGGAGTTGGAAACCGTTCTGAGTGGTGCTCT 2100
DB |||||
QY 2101 GAGTCTGAAGAGGAGCTAGCTGCGGGTTCTGCGACTCTTAGAGTCTCTTACTCTGTTTAGAA 2160
DB |||||
QY 2101 GAGTCTGAAGAGGAGCTAGCTGCGGGTTCTGCGACTCTTAGAGTCTCTTACTCTGTTTAGAA 2160
DB |||||
QY 2161 CTTTCAAGGGTACAAAGTGGAACTGGACTTAAAGAGTTTTTAAATGAATAGACTTCAT 2220
DB |||||
QY 2221 TCTGCTTTTGTGGCTTTGAGTTTTTAAAGTAACTTGTCTGGGCTATTCTCTGTTTATAGT 2280
DB |||||
QY 2221 TCTGCTTTTGTGGCTTTGAGTTTTTAAAGTAACTTGTCTGGGCTATTCTCTGTTTATAGT 2280
DB |||||
QY 2281 CACAATAATTTATAGAAATGAAGGTGTAAATTAAGTCTCTTTTATTAATTAATTCATAGCA 2340
DB |||||
QY 2281 CACAATAATTTATAGAAATGAAGGTGTAAATTAAGTCTCTTTTATTAATTCATAGCA 2340
DB |||||
QY 2341 TTATCAAACTCTCCAGGTAGCAAAACACACAGTCAATTTGATTTGGCGATTTACAAGCGAG 2400
DB |||||
QY 2341 TTATCAAACTCTCCAGGTAGCAAAACACACAGTCAATTTGATTTGGCGATTTACAAGCGAG 2400
DB |||||
QY 2401 AATATCTGGTTTTGATGATTTATATGTCTATTTACTGTAGCTTAAATATGTTTAAAA 2460
DB |||||
QY 2401 AATATCTGGTTTTGATGATTTATATGTCTATTTACTGTAGCTTAAATATGTTTAAAA 2460
DB |||||
QY 2461 TGACTTTGAGTGAATGTTTTGCGCAACCTAGGGGTTTATGGATCAGAAATGCGGCTGAG 2520
DB |||||
QY 2461 TGACTTTGAGTGAATGTTTTGCGCAACCTAGGGGTTTATGGATCAGAAATGCGGCTGAG 2520
DB |||||
QY 2521 TCGTTTGGTTTTGAGTCTAGAGTTTTTCAAGGGCAAAATCAAACTAGCTTGTTTCTG 2580
DB |||||
QY 2521 TCGTTTGGTTTTGAGTCTAGAGTTTTTCAAGGGCAAAATCAAACTAGCTTGTTTCTG 2580
DB |||||
QY 2581 GTTCACTTACCAGCTCGGGAGCTGCCCTTATATATTTCTACATAGGACTTATTATACA 2640
DB |||||
QY 2581 GTTCACTTACCAGCTCGGGAGCTGCCCTTATATATTTCTACATAGGACTTATTATACA 2640
DB |||||
QY 2641 TAAAGTCTGTATAAATGTCCTGAAGATGACACCTAGCTGCTTCACTCGAAGGGTCTCT 2700
DB |||||
QY 2641 TAAAGTCTGTATAAATGTCCTGAAGATGACACCTAGCTGCTTCACTCGAAGGGTCTCT 2700
DB |||||
QY 2701 GGGGCTGGAGAGTTGGTCAAGTTAAGAAACCGGTGTTGCTCTCTCCAGACGACCCAG 2760
DB |||||
QY 2701 GGGGCTGGAGAGTTGGTCAAGTTAAGAAACCGGTGTTGCTCTCTCCAGACGACCCAG 2760
DB |||||
QY 2761 CGACTTGTGAGGCTCATTAACCGACCACTCCAGCCTGGGGCATCTGACGCCCTCTCTGA 2820
DB |||||
QY 2761 CGACTTGTGAGGCTCATTAACCGACCACTCCAGCCTGGGGCATCTGACGCCCTCTCTGA 2820
DB |||||
QY 2821 CTTTAAAGGGCAATTTGGTGACACGTTACATACAGGAGGCAAAACATTTTATACATGAAC 2880
DB |||||
QY 2821 CTTTAAAGGGCAATTTGGTGACACGTTACATACAGGAGGCAAAACATTTTATACATGAAC 2880
DB |||||
QY 2881 GTAATAATGCAATAGTTAGTGAAGCGCTGAAGGAAAGAGTTTTAGATGCAAGGTTTA 2940
DB |||||
QY 2881 GTAATAATGCAATAGTTAGTGAAGCGCTGAAGGAAAGAGTTTTAGATGCAAGGTTTA 2940
DB |||||
QY 2941 GTCTGACCTGAGTGGCTTCCCTTAGCCGGATGGTGTCTCTCAAGTTCAGGAGAACCAT 3000
DB |||||
QY 2941 GTCTGACCTGAGTGGCTTCCCTTAGCCGGATGGTGTCTCTCAAGTTCAGGAGAACCAT 3000
DB |||||
QY 3001 TTCTTTTTTATCAGGAAAAAGGACTTATTCCTTAGGGCTCTGCTGACATCCCTTAGGAACA 3060
DB |||||
QY 3001 TTCTTTTTTATCAGGAAAAAGGACTTATTCCTTAGGGCTCTGCTGACATCCCTTAGGAACA 3060
DB |||||
QY 3061 GAGATAAAATACGGATGGATGGAATGAACATGCTTTGGGTGAAGGAGCGAGTACCTGA 3120
DB |||||
QY 3061 GAGATAAAATACGGATGGATGGAATGAACATGCTTTGGGTGAAGGAGCGAGTACCTGA 3120
DB |||||
QY 3121 CTGGACCCAGTGGGCCACTTCCACAAGCGAAGCCCGGTAAACCGATGTGCACCTCCAGAAATC 3180

DB |||||
QY 3121 CTGGACCCAGTGGGCCACTTCCACAAGCGAAGCCCGGTAAACCGATGTGCACCTCCAGAAATC 3180
DB |||||
QY 3181 TTCTCCCTTCTGGTAGAGATGACATCGATGAGTGTGTCTGACCACTGGGCGCTGCATG 3240
DB |||||
QY 3181 TTCTCCCTTCTGGTAGAGATGACATCGATGAGTGTGTCTGACCACTGGGCGCTGCATG 3240
DB |||||
QY 3241 GGTGTGGGCTCTGTTCTTCCGAATCTA CTTGAGATCTCAGGACAGAGGAAGCCATGA 3300
DB |||||
QY 3241 GGTGTGGGCTCTGTTCTTCCGAATCTA CTTGAGATCTCAGGACAGAGGAAGCCATGA 3300
DB |||||
QY 3301 ATGTTACCAAGTGGTCTGAGTGTGTGATGATTTTGAACCATTTATTTGGTTTTT 3360
DB |||||
QY 3301 ATGTTACCAAGTGGTCTGAGTGTGTGATGATTTTGAACCATTTATTTGGTTTTT 3360
DB |||||
QY 3361 AGGAGAAATTTCTCTCTCTAGTGCCCTGTGATGCCAAAGCCAGCCCTTTCAGAAAGTGTTC 3420
DB |||||
QY 3361 AGGAGAAATTTCTCTCTCTAGTGCCCTGTGATGCCAAAGCCAGCCCTTTCAGAAAGTGTTC 3420
DB |||||
QY 3421 CTTCTGTCTGCTTCCGTTTGTAAATGTGATCCACTCGGGGAAATGGGTGTCTACCCAGGA 3480
DB |||||
QY 3421 CTTCTGTCTGCTTCCGTTTGTAAATGTGATCCACTCGGGGAAATGGGTGTCTACCCAGGA 3480
DB |||||
QY 3481 AAGTGCCTACCCAGTCAAGGCTCAACAGTGTGTTGAGGATCAAAACATGGCTCTGT 3540
DB |||||
QY 3481 AAGTGCCTACCCAGTCAAGGCTCAACAGTGTGTTGAGGATCAAAACATGGCTCTGT 3540
DB |||||
QY 3541 GAAATACTGCCACCATTCTATTTGGTTGGATTTCTCAGGAGTCTGAAATCTTCCCTCAGC 3600
DB |||||
QY 3541 GAAATACTGCCACCATTCTATTTGGTTGGATTTCTCAGGAGTCTGAAATCTTCCCTCAGC 3600
DB |||||
QY 3601 AGTCTCTTCTCCAAACCCCTACCCAGAGCCAACTGGGATTTGAACTTCTCTGACTC 3660
DB |||||
QY 3601 AGTCTCTTCTCCAAACCCCTACCCAGAGCCAACTGGGATTTGAACTTCTCTGACTC 3660
DB |||||
QY 3661 TCTTCTTCCCTCAGGCTTGACAACTAAAGGTCTCTGGGACACCCAGCTAGGSCCTTCCC 3720
DB |||||
QY 3661 TCTTCTTCCCTCAGGCTTGACAACTAAAGGTCTCTGGGACACCCAGCTAGGSCCTTCCC 3720
DB |||||
QY 3721 CAACTCTTATCAGCTGAACTTGGATTTCTTCCCAACAGGCTTGAACCTGGGTGCTGTT 3780
DB |||||
QY 3721 CAACTCTTATCCAGCTGAACCTTGGATTTCTTCCCAACAGGCTTGAACCTGGGTGCTGTT 3780
DB |||||
QY 3781 GTTCCACTGGCCAAACAAACATCTTTTGGCCAGATTTGGATTTCTCAATAGATTTTATAGA 3840
DB |||||
QY 3781 GTTCCACTGGCCAAACAAACATCTTTTGGCCAGATTTGGATTTCTCAATAGATTTTATAGA 3840
DB |||||
QY 3841 CAATATTTCTCCACAGACTTTTAAACATGGCTTGTCTTTCCATACACATCCGGTCA 3900
DB |||||
QY 3841 CAATATTTCTCCACAGACTTTTAAACATGGCTTGTCTTTCCATACACATCCGGTCA 3900
DB |||||
QY 3901 TTTAAACATTTTATTAACCAAGGAAATTTAAACCAAGCAATAGAGTACTTTTCAGATATA 3960
DB |||||
QY 3901 TTTAAACATTTTATTAACCAAGGAAATTTAAACCAAGCAATAGAGTACTTTTCAGATATA 3960
DB |||||
QY 3961 AACTGTGTTTCACTTTTATGTAGAGTGTGCTATGTATAGGCGGTATGTACCCGTGCTGA 4020
DB |||||
QY 3961 AACTGTGTTTCACTTTTATGTAGAGTGTGCTATGTATAGGCGGTATGTACCCGTGCTGA 4020
DB |||||
QY 4021 AGTAATATTAACCATAGTCTCTGGGAGGATTTACAGACCTTTTTCGACTTTATGCTTTTTG 4080
DB |||||
QY 4021 AGTAATATTAACCATAGTCTCTGGGAGGATTTACAGACCTTTTTCGACTTTATGCTTTTTG 4080
DB |||||
QY 4081 TGAACCTCGATAACCATGGTCAATATTAAGCCAAATACTGGCAATTTCTGTGAATAAAC 4140
DB |||||
QY 4081 TGAACCTCGATAACCATGGTCAATATTAAGCCAAATACTGGCAATTTCTGTGAATAAAC 4140
DB |||||
QY 4141 ATGCAATATGATCTTAAAAAATAAAAAAAAAAAAA 4174
DB |||||
QY 4141 ATGCAATATGATCTAAAAAATAAAAAAAAAAAAA 4174
DB |||||

BC059819 4078 bp mRNA linear ROD 21-OCT-2003
LOCUS Mus musculus cancer related gene-liver 1, mRNA (cDNA clone
DEFINITION MGC:69583 IMAGE:6839525), complete cds.
ACCESSION BC059819
VERSION BC059819.1 GI:37590519
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4078)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Rodriguez, R.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED 12477932
2 (bases 1 to 4078)
AUTHORS Strausberg, R.
DIRECT SUBMISSION
SUBMITTED (07-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIB-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Neas, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smalusz, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Nataesja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: JRAK Plate: 132 Row: 0 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA GI: 21314857.
Location/Qualifiers
1. .4078
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:69583 IMAGE:6839525"
/tissue_type="Brain"

clone_lib="NIH_BMAP_GH0"
lab_host="DHI0B"
note="Vector: pYX-ASC"
1. .4078
/gene="CRG-L1"
/db_xref="LocusID:230379"
76. .765
/gene="CRG-L1"
/codon_start=1
/product="CRG-L1 protein"
/protein_id="AAH59819.1"
/db_xref="GI:37590520"
/db_xref="LocusID:230379"
/translation="MGAPHHMDHHRAGSEVDMCEDNTIVPAIAEFNTISNVLPFI
LAPICMLFRQYATCFNSGIYLIWTLVVGIGSVYFHALSFLGMDLALVLWM
CLAMWFFRRYLPKIFRNDRCNDRVFKLGLFSLGMLTALFQWISDAQCELLSSFH
PYYLHCVMHILICLASVLCVCFAYFDAASBIPEQGVIREWSEKNAFIVGPVYVSL
CAHKKSFPVKIT"

ORIGIN

Query Match 92.6%; Score 3864.4; DB 10; Length 4078;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 4025; Conservative 0; Mismatches 1; Indels 140; Gaps 2;

QY 10 GCTGCTCGATGCTCCAGAGCGGCATCGGGCGCCCGCACCTGGTGGGACCACTGCGGGC 69
DB 51 GCTGCTCGATGCTCCAGAGCGGCATCGGGCGCCCGCACCTGGTGGGACCACTGCGGGC 110
QY 70 TGGCAGTTTCGGAGGTGGATTGGTGGCGAGCAAACTACACTATCGTGCCTGCCATTGCCGA 129
DB 111 TGGCAGTTTCGGAGGTGGATTGGTGGCGAGCAAACTACACTATCGTGCCTGCCATTGCCGA 170
QY 130 GTTCTACACAGATCAGCAACGCTCTGTTTTTATTTTACCTCCCATCTGCATGTGCTT 189
DB 171 GTTCTACACAGATCAGCAACGCTCTGTTTTTATTTTACCTCCCATCTGCATGTGCTT 230
QY 190 GTTCCGCGAGTACCAACGCTCTTCAACAGCGGCATCTACTTAATATGAGCGCTCCTAGT 249
DB 231 GTTCCGCGAGTACCAACGCTCTTCAACAGCGGCATCTACTTAATATGAGCGCTCCTAGT 290
QY 250 TGTAGTGGGATTGGATCTGTCTACTTCCATCGCAACGCTGAGTTTCTTGGGTTCAGATGCT 309
DB 291 TGTAGTGGGATTGGATCTGTCTACTTCCATCGCAACGCTGAGTTTCTTGGGTTCAGATGCT 350
QY 310 TGATGAATTCGCCATTCTGTGGGTTCTGATGTGTGCTTTGGCCATGTGGTTTCCAGAG 369
DB 351 TGATGAATTCGCCATTCTGTGGGTTCTGATGTGTGCTTTGGCCATGTGGTTTCCAGAG 410
QY 370 GTATTTACCAAGATCTTTCGGAATGACAGGGCGAGGTTCAAGGCAGTGTGTGTCTCT 429
DB 411 GTATTTACCAAGATCTTTCGGAATGAC-----AGGTGTGAATGT 438
QY 430 GTCTGCAATTAACAAGCTGTTGGCGTTTATCAAGCCCGCCATCAACAATATTTCCCTGAT 489
DB 439 ----- 438
QY 490 GATTCTGGGACTTCCATGCACCTGCGCTGCTGTTGTCAGAGCTGAAGAGGTGTGACAAATGT 549
DB 439 -----AGGTGTGAATGT 452
QY 550 GCGTGTGTTTAAAGCTGGGCTCTTTCTCTGGGCTCTGGTGGACTCTGGCTCTCTTCTGCTG 609
DB 453 GCGTGTGTTTAAAGCTGGGCTCTTTCTCTGGGCTCTGGTGGACTCTGGCTCTCTTCTGCTG 512
QY 610 GATCAGGACCAAGCCCTTCTGTGAGCTGCTCTCTCTCTTCACTTCCCTTCACTTCCCTTCACTGCACTG 669
DB 513 GATCAGGACCAAGCCCTTCTGTGAGCTGCTCTCTCTCTTCACTTCCCTTCACTTCCCTTCACTGCACTG 572
QY 670 TGTGTGCATATTTCTCATCTGCCCTTGTCTTCTGCTGCTGCTGTGTGCTGTGCTGTGCTT 729
DB 573 TGTGTGCATATTTCTCATCTGCCCTTGTCTTCTGCTGCTGTGTGCTGTGCTGTGCTTCT 632
QY 730 TGATGCTGCTTCAGAGATACCTGAGCAAGGTCCTCAGTTCATCAGATTCTGGCCCGCAGGAA 789

QY 2948 CTGAGTGCCTTCTTAGCGCGATGCTGCTCTCAAGTCAGGAGAACCACTTCTTTT 3007
 Db 2853 CTGAGTGCCTTCTTAGCGCGATGCTGCTCTCAAGTCAGGAGAACCACTTCTTTT 2912
 QY 3008 TTATCAGGAAAAGGACTTATCCCTAGGCGCTCTGCTGACATCCCTAGGAAACAGAGATAA 3067
 Db 2913 TTATCAGGAAAAGGACTTATCCCTAGGCGCTCTGCTGACATCCCTAGGAAACAGAGATAA 2972
 QY 3068 AATACGGATGGATGTAATGAACATGCTTGGGTGAAGAGCGGAGTACTGCTGAGACC 3127
 Db 2973 AATACGGATGGATGTAATGAACATGCTTGGGTGAAGAGCGGAGTACTGCTGAGACC 3032
 QY 3128 CAGTGGGCGCCTTCCACAGCGAAGCCGCTTAACGATGTCGACATCCAGAACTTCTCCC 3187
 Db 3033 CAGTGGGCGCCTTCCACAGCGAAGCCGCTTAACGATGTCGACATCCAGAACTTCTCCC 3092
 QY 3188 TTCTGGTAGAGATGATCATGATGAGTGTGCTGACCACTGGGCGCTGATGGGTGCG 3247
 Db 3093 TTCTGGTAGAGATGATCATGATGAGTGTGCTGACCACTGGGCGCTGATGGGTGCG 3152
 QY 3248 GGCCTGTTCTGTCGGAATCTACCTGAGATCTCAGGACAGAGGAGCCATGAATGTAC 3307
 Db 3153 GGCCTGTTCTGTCGGAATCTACCTGAGATCTCAGGACAGAGGAGCCATGAATGTAC 3212
 QY 3308 CAAGTGGTCAATGGCTGTCAGTGAATTTTACAGTTTGAACCAATATTGTTGTTTTAGAGAA 3367
 Db 3213 CAAGTGGTCAATGGCTGTCAGTGAATTTTACAGTTTGAACCAATATTGTTGTTTTAGAGAA 3272
 QY 3368 TTCTTCTCTCTAGTGCCTGTGATGCAAGAGCCGCTTCAAGAGTGTTCCTTCTGT 3427
 Db 3273 TTCTTCTCTCTAGTGCCTGTGATGCAAGAGCCGCTTCAAGAGTGTTCCTTCTGT 3332
 QY 3428 CTGCTTCCCTTTGTAATGATGATCCACTCGGGAATGGTGTCTACCCAGGAAAAGTGC 3487
 Db 3333 CTGCTTCCCTTTGTAATGATGATCCACTCGGGAATGGTGTCTACCCAGGAAAAGTGC 3392
 QY 3488 CTACCCCACTCAGCGGTCAACAGTGTGTGTGAGGATCAACATGCTGTGTGAAAATA 3547
 Db 3393 CTACCCCACTCAGCGGTCAACAGTGTGTGTGAGGATCAACATGCTGTGTGAAAATA 3452
 QY 3548 CTGCAACCGATCTAATGGTGTGATCTCAGGAGTCTGAATCTTCCCTCAGAGTCTCTC 3607
 Db 3453 CTGCAACCGATCTAATGGTGTGATCTCAGGAGTCTGAATCTTCCCTCAGAGTCTCTC 3512
 QY 3608 TTCTCCCAACCCCTACCCAGAGCCCAACTCGGATTTGAACTTCTGACTCTCTTCTT 3667
 Db 3513 TTCTCCCAACCCCTACCCAGAGCCCAACTCGGATTTGAACTTCTGACTCTCTTCTT 3572
 QY 3668 CCTCAGGTCTGACAACTAATGGTCTCTGGGGACACCCAGCTAGGGCCTTCCCAACTCC 3727
 Db 3573 CCTCAGGTCTGACAACTAATGGTCTCTGGGGACACCCAGCTAGGGCCTTCCCAACTCC 3632
 QY 3728 TTATCCAGCTGAATCTTCCCAACAGGCGCTTGAAGTGGTCTGTGTGTCCTCA 3787
 Db 3633 TTATCCAGCTGAATCTTCCCAACAGGCGCTTGAAGTGGTCTGTGTGTCCTCA 3692
 QY 3788 CTGGCAACACACATCTTGGCCAGATGGGATTTCTCAATAGATTTTATAGACATTTT 3847
 Db 3693 CTGGCCACACACATCTTGGCCAGATGGGATTTCTCAATAGATTTTATAGACATTTT 3752
 QY 3848 CTCCCAACAGACTTTAAACATGGCTTGTGTCTTTCATACACATCCCGTCTGAGTTTAA 3907
 Db 3753 CTCCCAACAGACTTTAAACATGGCTTGTGTCTTTCATACACATCCCGTCTGAGTTTAA 3812
 QY 3908 CTATTTTATACCAACAGAAATTAACCAAGCAATAGAGTACTTTCAGATATAAAGTGTG 3967
 Db 3813 CTATTTTATACCAACAGAAATTAACCAAGCAATAGAGTACTTTCAGATATAAAGTGTG 3872
 QY 3968 TTTTCATCTTTATGTAGAGTGTGTCTATCTATAGGCGGTATCTACCTGGCTGAAGTAATA 4027
 Db 3873 TTTTCATCTTTATGTAGAGTGTGTCTATCTATAGGCGGTATCTACCTGGCTGAAGTAATA 3932

QY 4028 TTAAACATAGCTCTGGAGGATTTACAGACCTTTTGACCTTTATGCTTTTGTGAACTC 4087
 Db 3933 TTAAACATAGCTCTGGAGGATTTACAGACCTTTTGACCTTTATGCTTTTGTGAACTC 3992
 QY 4088 TGATAACCATGTCATATTTAAAGCCAATAATGCGCATTTTCTGTGAATAAATCATGCATA 4147
 Db 3993 TGATAACCATGTCATATTTAAAGCCAATAATGCGCATTTTCTGTGAATAAATCATGCATA 4052
 QY 4148 TGTATCTATAAAAAAAAAAAAAAAAAAAAA 4173
 Db 4053 TGTATCTATAAAAAAAAAAAAAAAAAAAAA 4078

RESULT 3
 EX005084
 LOCUS
 DEFINITION
 Mouse DNA sequence from clone RP24-468M3 on chromosome 4, complete sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 90370)
 Barlow,K.
 Direct Submission
 Submitted (22-OCT-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Oct 23, 2003 this sequence version replaced gi:37693442.
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP24-468M3 is from a Male (C57BL/6J) mouse BAC Library VECTOR: pTARBAC1.
 FEATURES
 source
 Location/Qualifiers
 1..90370
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clone="RP24-468M3"
 /clone_lib="RPC1-24"

Query Match 83.1%; Score 3467.6; DB 10; Length 90370;

	Best Local Similarity	99.5%; Pred. No. 0;	Matches	Conservative	0; Mismatches	14; Indels	2; Gaps	1;
QY	654	TCCCTACTGACGTGGTGGCATATTTCTCATCTGCCTTGCTTCGTACCTGGGCTGTG	713					
Db	78025	TCACATCTCCCTTCTCTGCGAGCAATTTCTCATCTGCCCTTGCTTCGACCTGGGCTGTG	78084					
QY	714	TGTGCTTGCCCTAATTGATGTGCTCAGAGATACCTGAGCAAGGTCCAGTCATCAGAT	773					
Db	78085	TGTGCTTGCCCTACTTTGATGTGCTCAGAGATACCTGAGCAAGGTCCAGTCATCAGAT	78144					
QY	774	TTGCGCCAGCGAGAATAATGGGCTTTTAATGTTGTGCCCTTAATGTGTCCCTTCTGTGTGCC	833					
Db	78145	TCTGGCCCCAGCGAGAAATGGGCTTTTAATGTTGTGCCCTTATATGTGTCCCTTCTGTGTGCC	78204					
QY	834	ACAAGAGTGCAGCTCAAGATCACTGATGCAAGGACGAGTGCACCAGCTTCTCTACTTTAC	893					
Db	78205	ACAAGAAGTGCAGCTCAAGATCACTGATGCAAGGACGAGTGCACCAGCTTCTCTACTTTAC	78264					
QY	894	TTCTATTTCAGTGCCTGGGCTTCGTTTGTCTAGCAAAGATGGGCTGAGGGGTTGAGGA	953					
Db	78265	TTCTATTTCAGTGCCTGGGCTTCGTTTGTCTAGCAAAGATGGGCTGAGGGGTTGAGGA	78324					
QY	954	ATTGTTGTGGGTGTTTAAAATTTCTGCTCCTTTGTGATCTAACTGGAACACTGTGC	1013					
Db	78325	ATTGTTGTGGGTGTTTAAAATTTCTGCTCCTTTGTGATCTAACTGGAACACTGTGC	78384					
QY	1014	CTGACCTCCTTAGTTAAGTAGGGCTCAGGACATAGGTTCTCTCTCAGTATCCTT	1073					
Db	78385	CTGACCTCCTTAGTTAAGTAGGGCTCAGGACATAGGTTCTCTCTCAGTATCCTT	78444					
QY	1074	TCCAGGCATACGGGCTTGCTGGGTTATGTGCCATATGATCAACAGAGTAGTTCCTTGG	1133					
Db	78445	TCCAGGCATACGGGCTTGCTGGGTTATGTGCCATATGATCAACAGAGTAGTTCCTTGG	78504					
QY	1134	GAGACCTAGGGCAACCCAAGTTTCTGTCTGGAGAGGGTAGCTTGCTTTCCACATTT	1193					
Db	78505	GAGACCTAGGGCAACCCAAGTTTCTGTCTGGAGAGGGTAGCTTGCTTTCCACATTT	78564					
QY	1194	CNAGCTCTCAACCCCTGATTAACAGACCTCTGTATTTGTTGATGAAGGTTCCAGAA	1253					
Db	78565	CCAGACTCTCAACCCCTGATTAACAGACCTCTGTATTTGTTGATGAAGGTTCCAGAA	78624					
QY	1254	CTTTTCATTTTGGCGGAGAACTGTCTTCAACAAAAACCAAGTGGGCAAAAACAGTGTG	1313					
Db	78625	CTTTTCATTTTGGCGGAGAACTGTCTTCAACAAAAACCAAGTGGGCAAAAACAGTGTG	78684					
QY	1314	GGGGTGTGACTAAGACGGGTGCTTGTCACTCGCACCTCTTAGCTTTCCAGGTTCTCC	1373					
Db	78685	GGGGTGTGACTAAGACGGGTGCTTGTCACTCGCACCTCTTAGCTTTCCAGGTTCTCC	78744					
QY	1374	ACGTGTTTGGATCCTTCCACTACCTCTGCTGAGAGTGGAGCCACGCTTCAGAGGC	1433					
Db	78745	ACGTGTTTGGATCCTTCCACTACCTCTGCTGAGAGTGGAGCCACGCTTCAGAGGC	78804					
QY	1434	AAAGCTGGCAACCCCTCTATGCAAAAAGTACACTCTCTTTAAGCACACATTAAC--AT	1491					
Db	78805	AAAGCTGGCAACCCCTCTATGCAAAAAGTACACTCTCTTTAAGCACACATTAAT	78864					
QY	1492	AGACACTATTTCTGCTCTTCCAGAGTGACAGCCCTCAGACCCACAGAGAATCCTTFCAGG	1551					
Db	78865	AGACACTATTTCTGCTCTTCCAGAGTGACAGCCCTCAGACCCACAGAGAATCCTTFCAGG	78924					
QY	1552	TTATGTGAAGATTCCACACACAGCCCCCTTGTGACCTCTGTGAGAGCAAGCCCTCGTT	1611					
Db	78925	TTATGTGAAGATTCCACACACAGCCCCCTTGTGACCTCTGTGAGAGCAAGCCCTCGTT	78984					
QY	1612	GTAATCGCGAGGACGCTCTGGAGGCTTGACTGTGGAGGCCCTGTGAAGATTTTGAAGAG	1671					
Db	78985	GTAATCGCGAGGACGCTCTGGAGGCTTGACTGTGGAGGCCCTGTGAAGATTTTGAAGAG	79044					
QY	1672	CTTTTCTTATTTGGAAGCTTTTGGTGTGCTGTACCTCCAGAATTTCTGA CCCCCTTTCTCT	1731					
Db								

[illegible]

QY	2332	TTCTATGACATTTTACCAACTCTCCAGGTAGCAAAACACAGATCATTTGATTTGGGATTTA	2391
Db	43911	TTCTATGACATTTTACCAACTCTCCAGGTAGCAAAACACAGATCATTTGATTTGGGATTTA	43970
QY	2392	CAAGGACGAATACTTGGTTTGAATGATGTATATGTATGTCATTTTACTGTAGCTTAAATA	2451
Db	43971	CAAGGACGAATACTTGGTTTGAATGATGTATATGTATGTCATTTTACTGTAGCTTAAATA	44030
QY	2452	TGTTTAAATGACTTTGAGTGAATCTTTGTGGCAACCTAGGGTTTATGGATCAGAAAT	2511
Db	44031	TGTTTAAATGACTTTGAGTGAATCTTTGTGGCAACCTAGGGTTTATGGATCAGAAAT	44090
QY	2512	GGGGCTGAGTCCTTTGGTTTGGAGTCTAGAAGTTTTCAGAGGGCAAAATCAAACTAGCC	2571
Db	44091	GGGGCTGAGTCCTTTGGTTTGGAGTCTAGAAGTTTTCAGAGGGCAAAATCAAACTAGCC	44150
QY	2572	TTGTTTCTGGTTCATCTTACAGGCTCGGAGCTGCCCTTATATATTTCTACATAAGGACT	2631
Db	44151	TTGTTTCTGGTTCATCTTACAGGCTCGGAGCTGCCCTTATATATTTCTACATAAGGACT	44210
QY	2632	TATTTATACATAAGTCTGTATATAATGCTTGAAGATGACACCTAGCTGCCCTTCTGCGAA	2691
Db	44211	TATTTATACATAAGTCTGTATATAATGCTTGAAGATGACACCTAGCTGCCCTTCTGCGAA	44270
QY	2692	GGGTCTCTGGGGCTGGAGGTTGGTTCAACAGTTTAAAGAACCCGCTGTCTCTCCAG	2751
Db	44271	GGGTCTCTGGGGCTGGAGGTTGGTTCAACAGTTTAAAGAACCCGCTGTCTCTCCAG	44330
QY	2752	AGGACCCAGCGACTTGTGAGGCTCATACACAGAACTCCAGCCCTGGGGCATCTGACGCC	2811
Db	44331	AGGACCCAGCGACTTGTGAGGCTCATACACAGAACTCCAGCCCTGGGGCATCTGACGCC	44390
QY	2812	CTCTCTGACTTCTAAGGSCATTTGGTGACACGTTACATACAGGCGAGCAAAACATTTAT	2871
Db	44391	CTCTCTGACTTCTAAGGSCATTTGGTGACACGTTACATACAGGCGAGCAAAACATTTAT	44450
QY	2872	ACATGTAAACGTAAATAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	2931
Db	44451	ACATGTAAACGTAAATAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	44510
QY	2932	CAAGGTTTAGTCTGACCTGAGTGCCCTTCTAGCCGGATGGTGCTCTCAAGTCAAGG	2991
Db	44511	CAAGGTTTAGTCTGACCTGAGTGCCCTTCTAGCCGGATGGTGCTCTCAAGTCAAGG	44570
QY	2992	AAGAACCAATTTCTTTTATCAGGAABAGGACTTATTCCTTAGGCGCTCTGCTGACATCC	3051
Db	44571	AAGAACCAATTTCTTTTATCAGGAABAGGACTTATTCCTTAGGCGCTCTGCTGACATCC	44630
QY	3052	CTAGGAACAGAGATAAAATACGGATGGATGTGAATGAACATGCTTGGGTGAAGAGCGCG	3111
Db	44631	CTAGGAACAGAGATAAAATACGGATGGATGTGAATGAACATGCTTGGGTGAAGAGCGCG	44690
QY	3112	AGTACCTGACCTGACCCAGTGGGCCACTTCCACAGGAGCCCGGTAAACCGATGTGCAC	3171
Db	44691	AGTACCTGACCTGACCCAGTGGGCCACTTCCACAGGAGCCCGGTAAACCGATGTGCAC	44750
QY	3172	TCCAGAAATCTTCTCCCTTCTCGTAGAGATGACATCGATGATGTGTCGTCACACTGGG	3231
Db	44751	TCCAGAAATCTTCTCCCTTCTCGTAGAGATGACATCGATGATGTGTCGTCACACTGGG	44810
QY	3232	CCCTGCAATGGGTGCGGGCTCTGTTCTGTTCCGAACTCTACCTGAGATCTCAGGACAGAG	3291
Db	44811	CCCTGCAATGGGTGCGGGCTCTGTTCTGTTCCGAACTCTACCTGAGATCTCAGGACAGAG	44870
QY	3292	AAGCCATGAATGTTACCAAGTGTCATGGCTGTGAGTATTTACAGTTTGAACCAATTA	3351
Db	44871	AAGCCATGAATGTTACCAAGTGTCATGGCTGTGAGTATTTACAGTTTGAACCAATTA	44930
QY	3352	TTGGTTTTTGGAGAAATCTCTCTCTAGTGCCCTGTGATGCAAGGCGAGCCCTTCA	3411
Db	44931	TTGGTTTTTGGAGAAATCTCTCTCTAGTGCCCTGTGATGCAAGGCGAGCCCTTCA	44990
QY	3412	GAAGTGTTCCTTCTGTGTGCTTCCCTTTGTAAATGTGATCCACTCGGGGAAATGGGTGTCT	3471
Db	44991	GAAGTGTTCCTTCTGTGTGCTTCCCTTTGTAAATGTGATCCACTCGGGGAAATGGGTGTCT	45050
QY	3472	ACCAGGGGAAAGTGGCTACCCAGTCAGGGTCAACAGTGTGTGTGTGGAGATCAACA	3531
Db	45051	ACCAGGGGAAAGTGGCTACCCAGTCAGGGTCAACAGTGTGTGTGTGGAGATCAACA	45110
QY	3532	TGGCTCTGTGAAATACTGCCACCCATTTCTATTTGGTTGGATTTCTCAGAGATCTGAATCT	3591
Db	45111	TGGCTCTGTGAAATACTGCCACCCATTTCTATTTGGTTGGATTTCTCAGAGATCTGAATCT	45170
QY	3592	TCCTCAGAGTCTCTTCTTCCCTCCACCCCTTCCAGAGGCAACACTCGGATTTGAACCTT	3651
Db	45171	TCCTCAGAGTCTCTTCTTCCCTCCACCCCTTCCAGAGGCAACACTCGGATTTGAACCTT	45230
QY	3652	CTCTGACTCTCTTCTTCCCTCAGGCTGTGACAACTAATGTCTCTGGGACACCCAGCTAG	3711
Db	45231	CTCTGACTCTCTTCTTCCCTCAGGCTGTGACAACTAATGTCTCTGGGACACCCAGCTAG	45290
QY	3712	GGCTTCCCAACTCTTATCCAGCTGAACCTTGGATTTCTTCCCAACAGGGCTTGAACCTT	3771
Db	45291	GGCTTCCCAACTCTTATCCAGCTGAACCTTGGATTTCTTCCCAACAGGGCTTGAACCTT	45350
QY	3772	GGTGCTGTGGTCCCACTGGCCCAACACACATCTTTGGCCAGATTTGGGATTTCTCAATAGA	3831
Db	45351	GGTGCTGTGGTCCCACTGGCCCAACACACATCTTTGGCCAGATTTGGGATTTCTCAATAGA	45410
QY	3832	TTTTATAGACATTTATTTCTCCACAGACTTTTAAACATGCTTGTCTTCTTCCATACACAT	3891
Db	45411	TTTTATAGACATTTATTTCTCCACAGACTTTTAAACATGCTTGTCTTCTTCCATACACAT	45470
QY	3892	CGGTGAGATTTAAACATTTTATTAACACAGAGATTTAAACCAAGCAATAGAGTACTT	3951
Db	45471	CGGTGAGATTTAAACATTTTATTAACACAGAGATTTAAACCAAGCAATAGAGTACTT	45530
QY	3952	TCAGATATAAATCTGTTTTCATCTTTATGTAGAGTGTCTATGTATAGGCGGTATGTAC	4011
Db	45531	TCAGATATAAATCTGTTTTCATCTTTATGTAGAGTGTCTATGTATAGGCGGTATGTAC	45590
QY	4012	CCTGGCTGAAGTAAATTTAAACATGCTCTGGGAGGATTTACAGACCTTTTGCATTTAT	4071
Db	45591	CCTGGCTGAAGTAAATTTAAACATGCTCTGGGAGGATTTACAGACCTTTTGCATTTAT	45650
QY	4072	GCTTTTGTGAACTCTGATTAACCATGTCTCAATATTTAAAGCAATACTGGCATTTTCTG	4131
Db	45651	GCTTTTGTGAACTCTGATTAACCATGTCTCAATATTTAAAGCAATACTGGCATTTTCTG	45710
QY	4132	TGAATAACATGCATATGTATCTA	4155
Db	45711	TGAATAACATGCATATGTATCTA	45734
RESULT 5			
AC120245			
LOCUS			
DEFINITION	Rattus norvegicus clone CH230-259012, WORKING DRAFT SEQUENCE, 5		
ACCESSION	AC120245		
VERSION	AC120245.4		
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 184163)		
AUTHORS	Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Aisbrooke, S., Amin, A., Anguliano, D., Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,		

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joliviet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kovis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheva, L., Louised, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nair, L.,
Nwakoelam, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, R.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlezyk, R., Wood, H., Wright, D.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Unpublished
Worley, K.C.

2 (bases 1 to 184163)

Direct Submission

Submitted (05-MAY-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 184163)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:22856351.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVT

Center clone name: CH230-259012

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 175276 bases at least Q40

Consensus quality: 176285 bases at least Q30

Consensus quality: 177113 bases at least Q20

Estimated insert size: 175972; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 5 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 177885: contig of 177885 bp in length

* 177886 177885: gap of unknown length

* 177986 179377: contig of 1392 bp in length

* 179378 179477: gap of unknown length

* 179478 180571: contig of 1094 bp in length

* 180572 180571: gap of unknown length

* 180672 182336: contig of 1565 bp in length

* 182337 182336: gap of unknown length

* 182337 184163: contig of 1827 bp in length.

* Location/Qualifiers

1..184163

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-259012"

1..1513

/note="wgs end extension

clone_end:Sp6"

complement(4969..5842)

/note="clone boundary

clone_end:Sp6

site:

end_sequence:BZ169793"

40368..40445

/note="clone boundary

clone_end:T7

site:

end_sequence:BZ169790"

171006..172144

/note="wgs end extension

clone_end:T7"

173214..174400

/note="wgs end extension

clone_end:T7"

175979..177885

/note="wgs end extension

clone_end:T7"

ORIGIN

Query Match 29.3%; Score 1222.4; DB 2; Length 184163;

Best Local Similarity 77.3%; Pred. No. 0;

Matches 1886; Conservative 0; Mismatches 351; Indels 203; Gaps 25;

QY 632 GAGCTCTCTCCCTTTCACCTTCCCTACCTGCGCATATTCATCTGC 691

12892 GAGAGCATGCGACACCGCTCACCTCCCTCCCTTGCAGGCATCTCATCTGC 12951

QY 692 CTTCGCTTCGTACCTGGGCTGTGTGCTTCGCTACTTTGATGCTGCCTCAGAGATACCT 751

12952 CTTCGCGCGTACTCTGGGCTGGTGTGCTTCGCTACTTCGATGCTGCCTCAGAGATACCT 13011

QY 752 GAGCAAGTCCAGTTCATCAGATTCTGGCCACGAGAAATGGGCTTTTATTGGTGTCCCT 811
Db 13012 GACCAAGTCCAGTTCATCAGATTCTGGCCACGAGAAATGGGCTTTTATTGGTGTCCCT 13071
QY 812 TATGTGTCCTCTCTGTGTGCCCAAGAAGTCCGACGCAAGATCAGCGTATGGCAAGGC 871
Db 13072 TATGTGTCCTCTCTGTGTGCCCAAGAAGTCCGACGCAAGATCAGCGTATGGCAAGGC 13131
QY 872 AGTGACCAAGTTCCTCTACTTACTTCTATTCGAGTGGCGCTGGGCTTCGTTTGTAGCAA 931
Db 13132 AGTGACCGGCTTTCTCTCTACTTACTTCTATTCGAGTGGCTTCGTTTGTAGCAA 13191
QY 932 AGATGGCTGAGGGGTTGAGGAATTTGATGTAGTGGGTGTTTAAATTCCTCTCTTTG 991
Db 13192 AGATGGCCAGGGGTTTGAAGAAATTTGATGTAGTGGGTGTTTAAATTCCTCTTTG 13251
QY 992 TGATCTAACTGGACCACTGTGCTGACCTCCCTAGGTTAAAGTA ---GAGGGCTCAGGGAC 1048
Db 13252 TGATCTAACTGGACCGCTGTGCTGACATCCCAAGGGTGGTAAATGGAGGGCTCAGGGAC 13311
QY 1049 ATAG-GTGTCTTCTCTAGTATCTTTCCAGGCATACGGGCTTGCTGGGTTATGTCCATA 1107
Db 13312 ATGAGTGGGTGGGCTCAGCGACCTTTTACGGGCATATCGGGCTTGCTGTGTATATCCATG 13371
QY 1108 ATGACATCAACAGAGTAGTTCCTTGGGAGACCTAGGGCAACCCAAAGTTTCTTGCTGGAG 1167
Db 13372 ATGACATTTGACAGAGTAGTATTTGGGAGACCCAGAGACGCCAAGCTTCTTGCTGGAA 13431
QY 1168 AGGTPAGCTTG- ---TGTTTCCACATTTCCAGACTCTCAACCCCTGATAAACAAGACC 1224
Db 13432 AAGTGGCTTGCCATAGCCCCACCATTTCAGGCTCTCTCATTAACAAGAGTCAAGCCC 13491
QY 1225 TTCTGATTTGGTGATGAAGTTTCAGAACTTTTCAATTTGCGGGAGAACTGTCTCTTC 1284
Db 13492 C-CTGATTCTGTGATAGAAAGTTCCAGAACTTTTCAATTTGCTGGGAGAACTGTCTCTTC 13550
QY 1285 AACAAAACCAAGTGGGCAAA --- 1304
Db 13551 AACAAATCGAGTAGGCAAAAACCAAAAACCAAAAACCAAAAACCAAAAACCAAAA 13610
QY 1305 --- 13611 AACAAAACCAAAAACCAAAAACCAAAAACCAAAAACCAAAAACCAAAAACCAAAA 13670
QY 1316 GGGTGTGACTAAGACGGTGGCTTGTCTCACTCGCACTCTAGCTTTCCAGGTTCTCCAC 1375
Db 13671 GGGTGTGACTAAGCCCGTGGCTTGTCTCACTTGCACCTCTTAGCT ---TTCTCCAC 13722
QY 1376 GTGTTTGTGGATCCTTCCACTACCTCTGCTGAGAGATGGAGCCACGGCTTCAGAGGGCAA 1435
Db 13723 ATGCTGTGGATCCTTCCACTACCTCTGCTGAGAGACGGAGCCTTGGCTTCAGAGGGCAA 13782
QY 1436 AGCTGGCAACACCTCTATGCAAAAAGTCACTCTCTTTAAGCACACATT ---ACATAG 1493
Db 13783 AGCCAGTAACACCTCTACGCAAGAGCTACAGTCCACTTTAAGCACACATTAACAACCTG 13842
QY 1494 ACATATTTCTGCTCTTCCAGAGTGCAGACCTCAGACCCACAGAGATCCTTCAGGTT 1553
Db 13843 ACATATGTCTGCTCTTCCAGAGTGCAGACCTCAGACCCACAGAGGCTCTCTG --- 13900
QY 1554 ATGTGAAGATTCCAACACAGCCCTCT --- 1604
Db 13901 -TGTGGCAATTCCACACATAGCCCTTACCGCTTGGCACTTGACCTCTGTGAGAGCAAGG 13959
QY 1605 CTTGTTGTAATCCGGAGGAGCTCTG-GAGGCTTGACTGTGGAGCCCTGTGAAGATT 1663
Db 13960 ACTCATTTGCATCTAGGAGAGCTCTGAGGGTTTGTGATTTAGAGCTCTCTGTGAAGCTT 14019
QY 1664 TTGG-AAAGCTTTTCTTATTGGAAGCTTTTGGTGTGCTTACCTCCAGATTCAGACC 1722
Db 14020 TTGGAAGCTTTTCTTATTGGAAGCTTACAGTGTGTTTATCTTACCTCCAGACTCTGACCC 14079

QY 1723 CGTTTCTCTGCTGTTTATTAGGGGATATGTGCTTCTGAGTGACGTAGCCTCCCTAGGATG 1782
Db 14080 CATTTTCTCTGCTGTTTTCAGGGGATGTGTGCTTCTGAGGGATGCCACTTCCCTGGGATG 14139
QY 1783 TGGGCTCCGGCTTGTGTTTTCATAATACCTGTGTCTAACTAGTGTCTCAGAGCACTTTG 1842
Db 14140 TGGACCTTCAGGCTTATTTTGTATATACC-GGTGCTAACTAGTGTCTCAGAGCCCTTTG 14198
QY 1843 CTCCTTTGATGTCTGGGCGGTCACTACACTCTGATTTGGCTGGCT--GGGATCCCCAGGAG 1901
Db 14199 CTCCTTTGATGTCTGGCCAGTCACGACACTCTGATTTGGCTGGCTAGGGGTCACAGGTAG 14258
QY 1902 GAAGGGCAGACTCTGACATGATGTCTCTCACCTGCACTCTACTGTCTTCTACTGGGCTC 1961
Db 14259 GAAGGGCAGACTCTGACCTGAATGTCTCTCACTGCACTCTGCTCTCTCACTGGGCTA 14318
QY 1962 TCTTCAGGGTATGAATGGGTATGGG-GTATCTCAGGGATGTGTTGTAATCAGGACCTT 2020
Db 14319 TCTCAGGGTATGAGTGGGAATTTGGGTATATCTCAGGGATGTGTTGTAATCAGGACCTT 14378
QY 2021 CTGCTTTCTGACATTTCCATTTGTCTAGTGTGTGAGACTCACTCAGGGGACTGGATTTGGAAC 2080
Db 14379 CTGCTTTCTGACATTTCCATTTGT-TGGGGTGGGACAGCGCTCAGGGGACTGCAAGTTGCTAC 14437
QY 2081 ACCGTTCTGAGTGTGTCTGAGTCTGAAGAGCTAGCTGCGGTTCTGSCACTTCTTAGG 2140
Db 14438 ACTGTTCTGAGTGGGCTCTGAGTTTGAAGAGTTGGC---CGGTTCTGGCGCTTCCAGG 14494
QY 2141 ATCTCTTACTCTGTTTGAACCTTCACAGGGTCAAAAGTGGGAACTGACCTTAAAGAGTT 2200
Db 14495 AACTCTTACTCCGTATAGAAGCTTCACAGGGTCCAAAGTGGAACTGACT--AGGAGTT 14552
QY 2201 TTTAAATGAATPAGACTTCATTCGCTTTTGTGGCTTTGGAGTTTAAAGTAACCTTGTCTG 2260
Db 14553 TTTAAATGCATAGGCTTCATTCGCTTCTGGGCTTTGAAGTTTGAAGTAGCTTGTCTA 14612
QY 2261 GGCTATTTCTGTTTATAGTCAAAATATTTATAGAACATGAAGTGTAAATAGTTGT 2320
Db 14613 GGCTCTTCTTGTCTGTAGTCAAAATATCTGTAGAACCCGAAATGTGTGAAGAAAGTCAAT 14672
QY 2321 CTTTATTAAATTCATAGCAATTTACCACTCTCCAGGTAGCAACACACAGCT---AT 2376
Db 14673 CTTTATTAAAGTTTATAGCAATTTACTAAATCTCCAGGTAGCAAGCCCATAGTCAATTTAT 14732
QY 2377 TTGATTTGGCAATTTAAGGAGGAGAACTTCTGTTTGTAAATGATTTGATATGTCAATTTA 2436
Db 14733 TTGATTTGGTGAATTTAAGGAGAGACTTGGGCTTTGAATGACTGTATATGTTATTTA 14792
QY 2437 CTGTAGCTTAAATATATCTTTAAATGACTTTGAGTGAATGTTTGTGGCAACCTA--CGG 2494
Db 14793 TTGCAGCTTAAATATATTTAAATGATTTTGAAGTTTAAATGTTTGTGGCAACCTAGGGG 14852
QY 2495 GTTTATGGATCAGAAATTTGCGGCTGAGTCTTGTGTTTGTAGTCTAGAAGTTTTCAGAGG 2554
Db 14853 GTTCATGGACAGAAATTTGCTGAGTCTTGTGTTTGTAGTCTAGAAATTTTTCAG--- 14909
QY 2555 GCAAAATCAAACTAGCCTTGTCTGTTTCACTTACAGGCTCGGAGCTGCGGCTTATA 2614
Db 14910 -----ACTAGCCTTGTCTTCCAGTCACTATCTCAGCAGGCAAGGGGCTGCTCTATA 14959
QY 2615 TATTCTACATAAGGACTTATTATACATAAGTCTGTATAAATGCTCTGAAGATGACACTA 2674
Db 14960 TATTCTACA-----CGATATACACATCTGAGGATGACGGCA 14997
QY 2675 GCTGCTTCTATCTGGAAAGGCTC-----GTCTGGGCTGGAGAGTTTGGTTCAACAGTTAA 2728
Db 14998 GCTGCTTCTCTGAAAGGCCCCGAGCTGTCTGGGCTGGAGAGCTGTTTTCAGCCGATAA 15057
QY 2729 GAACCCGTGTTGCTCTCTCCAG-----ACGACCCAGGACTTGTGAGGCTCAT 2777
Db 15058 AAACACATGTTGGCCTCCAGAGACCCAGAGTTTGAATTTCCAGCATCTTATAAAGCTGAC 15117
QY 2778 AACCAAGCAACTCCAGCCTCGGGCATCTGACGCCCTCTTCTGACTTCTAAGG----- 2829

TITLE	Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
JOURNAL	Weinstock,G. and Gibbs,R.A.
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 249734)
TITLE	Worley,K.C.
JOURNAL	Submitted
	Submitted (17-OCT-2001) Human Genome Sequencing Center, Department
	of Molecular and Human Genetics, Baylor College of Medicine, One
	Baylor Plaza, Houston, TX 77030, USA

AUTHORS
 TITLE
 JOURNAL
 COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly. In a contig scaffold, within each contig scaffold

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

shotgun sequence only contigs will be indicated in the feature table.

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center address: BCM

```

Center clone name: CH230-3E15
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 241576 bases at least Q40
Consensus quality: 243506 bases at least Q30

```

 * NOTE: Estimated insert size may differ from sequence length

- * [see http://www.ngsc.bcm.tmc.edu/docs/genbank_draft_data.html](http://www.ngsc.bcm.tmc.edu/docs/genbank_draft_data.html)
- * NOTE: This sequence may represent more than one clone.
- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 2 contigs. The true order of the pieces
- * is not known and their order in this sequence record is

* runs of N , but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1. 240455. Position of 240455 has been located.

* 248466	248565:	gap of unknown length
* 248566	249734:	contig of 1169 bp in length.
Location/Qualifiers		
1..249734		
/organism="Rattus norvegicus"		
FEATURES		
source		

```
misc_feature
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-3E15"
    1..2605
    /note="wds contig"
```

Query Match 29.3%; Score 1222.4; DB 2; Length 249734;
 Best Local Similarity 77.3%; Pred. No. 0;
 Matches 1886; Conservative 0; Mismatches 351; Indels 203; Gaps 25;

QY	632	GAGTGTCTCTCTCTTTTACATTTCCCTACCTGCACTGTGTGGCATATTTCTCATCTGC	691	Db	149955	ATCATTTGACACTCACTGAGCGAGCTCTGAGGGTTTTGATTTAGAGCTCTCTGTGAGCTT	150014
Db	148887	GAGAAAGCACTGACAGACCCCGCTCACCCCTCCCTCCCTTGGAGGCACATCTCATCTGC	148946	QY	1664	TTGG-AAAGCTTTTCTTATTTGGAAGCTTTTGTGTGTCCTTACCTCCAAGATTTCTGACCC	1722
QY	692	CTTGCTTCGTACTCGGGCTGTGTGTCTTTCGGCTACTTTGATGCTGCTCCCTCAGAGATACCT	751	Db	150015	TTGGAAGAGCTTTCTTATTTGGAAGCTTACAGTGTGTGTACCTCCAAGACTCTGACCC	150074
Db	148947	CTTGGCGGTACTCGGGCTGGGTGTGTCTTCGGCTACTTCGATGCTGCCTCAGAGATACCC	149006	QY	1723	CGTTCTCTGCTGTTTTTAAAGGGATATGTGCTTCCCTGAGTGTGAGTGTGCTCCCTGAGAG	1782
QY	752	GAGCAAGTCCAGTCAATCAGATTTCTGGCCAGGAGAAATGGGCTTTTATTTGGTGTCCCT	811	Db	150075	CATTTCTCTGCTGTTTTTACGGGGATGTGTGCTTCTGAGGGATGCCACTTCCCTCCGGATG	150134
Db	149007	GAAACAGGTCCAGTCAATCAGATTTCTGGCCAGTGTGAAATGGGCTTTTATTCGGTGTCCCC	149066	QY	1783	TTGGGCTCCGGCTTTTGTTTTCAATAACTCTGGTCTAACTGTGTTTTCTCAGAGCACTTTG	1842
QY	812	TATGTGTCCCTTCTGTGTGCCACAAAGAGTGCACAGTCAAGATCAAGTGTGGAAGGC	871	Db	150135	TGGAAGCTCCAGCTTTATTTTGTAAATACC-GGTGCTAACTGTGTTTTCTCAGAGCCCTTG	150193
Db	149067	TATGTGTCCCTCTGTGTGCCACAAAGAGTGCACAGTCAAGATCAAGTGTGGAAGGC	149126	QY	1843	CTCTCTTGATGCTGGGGCTCACTACACTCTGATTTGGTGGCT-GGGGATCCAGGGAG	1901
QY	872	AGTGACCAAGCTTCTCTACTTCTACTTTCTATTTGAGTGGCGCTGGGCTTCGTTTCTAGCAA	931	Db	150194	CTCTTCTGATGCTGCCAGTCAACACACTCTGATTTGGCTGGCTAGGGGTCAACAGTAG	150253
Db	149127	AGTGACCGGCTTTTCTGCTTACTTCTATTTGAGTGTGCTCTGGGCTTTGTTGCTAGAAA	149186	QY	1902	GAGGGGAGACTCTGACATGAATGTCTCTCACTGCACTCTACTGTCTTCACTGGCTC	1961
QY	932	AGATGGCTGAGGGGTGAGGAATTTGGTGTGTGGGTGTTTAAATTTCTGCTCCTTTG	991	Db	150254	GAGGGGAGACTCTGACATGAATGTCTCTCACTGCACTCTGCTCTCTCACTGGGCTA	150313
Db	149187	AGATGGCCAAAGGGTTTGAAGAAATGATGTAGTGTGGGTGTTTAAATTTCTGTTCTTTG	149246	QY	1962	TCTTCAGGGTATGAAGTGGGTATGGG-GTATCTCTCAGGGATGTTTTGTAACTCAGGCACTT	2020
QY	992	TGATCTAACTGGACCACTGTGCTGACCTCCCTAGGTAAAGTA--GAGGGCTCAGGGAC	1048	Db	150314	TCTGCAGGGTATGAGTGGAAATTTGGGTATATCTCAGGGATGTTTGTAACTCAGGCACTT	150373
Db	149247	TGATCTAACTGGACCGCTGTGCTGACATCCCAAGTGGGTGTTAAAGAGGGCTCAGGGAC	149306	QY	2021	CTGCTTTCTGACATTCATTTGCTAGTGTGTGAGACTGTCACTCAGGGGACTGGAGTTGGAAC	2080
QY	1049	ATAAG-GTGTCTTCTCAGTATCTTTTCCAGGCAATACGGGCTGCTGGGTATATGCCATA	1107	Db	150374	CTGCTTTCTGACATTCATTTGT-TGGGGTGGGACAGCGCTCAGGGGACTGCACTTTGCTAC	150432
Db	149307	ATGAGTGGGTGCGGCTCAGCGACTTTTACGGGCACATGGGCTGTGCTGTATATCCATG	149366	QY	2081	ACCGTTCTGAGTGTGCTCTGAGTCTGAAGGAGTCTGCTGCGGTTCTCGGCACTTCTAGG	2140
QY	1108	ATGACATCAACAGAGTATCTTTTGGAGACCTTAGGGCAACCCAAAGTTTCTTGTGGAG	1167	Db	150433	ACTGTTCTGAGTGGGCTCTGAGTTTGAAGAGTTGGC--CGTTCGCGCGTTCAGG	150489
Db	149367	ATGACATTCACAGAGTAGTTATTTGGGAGCCAGAGCGAGCCAAAGCTTCTTGTGTGAA	149426	QY	2141	ATCTCTTACTCTGTTTAGAACCTTCA CAGGGTACAAAGTGGGAACTGGACTTAAAGAGTT	2200
QY	1168	AGGCTAGCTTGC---TGTTTCCACCATTTCCAGACTCTCAACCCCTGATAAACAGACC	1224	Db	150490	AACTCTTACTCTCGTATAGAAGCTTCA CAGGGTCCAAAGTGGAACTGGACT--AGGAGTT	150547
Db	149427	AGGTTGGCTTGCATAGCCCCCACCATTTCAGGCTCTCTCATTACACAGGAGTCAAGCC	149486	QY	2201	TTTAAATGAATAGACTTCACTCTGCTTTGTGGCTTTGAGGTTTAAAGATTAAGTTGCTG	2260
QY	1225	TTCTGATTTGGTATGAAGGTTCCAGAACTTTTCATTTTGGCGGAGAGAACTGTGCTTTC	1284	Db	150548	TTTAAATGATAGGCTTCACTCTGCTCTGGGGCTTTTGAAGTTTAAAGATGAGTTGCTA	150607
Db	149487	C-CTGATTTCTGATAGAGGTTCCAGAACTTTTCATTTTGTGGGAGAAACTGTCTCTTC	149545	QY	2261	GGCTATTTCTGTTTATAGTCAAAATATTTATAGAAATGAAGGTGTAAATAAGTTGT	2320
QY	1285	AAACAAAACCAAGTGGCABA-----	1304	Db	150608	GGCTCTTCTGCTGTGTAGTCAAAATATCTGTAGAACCCGAATGTGTGAAGAGTCAAT	150667
Db	149546	AAACAAATCGATAGGCAAAAAA-----	149605	QY	2321	CTTTTATTAATTCATAGCAATTTACCACTCTCCAGGTAGCAAAACACAGCTC-----AT	2376
QY	1305	-----	1315	Db	150668	CTTTTATTAAGTTTATAGCATTTACTAAATCTCCAGGTAGCAAGCCCATAGTCAATTA	150727
Db	149606	AAACAAAAA-----	149665	QY	2377	TTGATTTGGGATTTACAGGAGGAAATACTTGGTTTGAATGATGATATATGTCATTTTA	2436
QY	1316	GGGTGTGACTAGACGGGTGGCTTTGTCTATCGACCTCTTAGCTTTTCCAGGTTCTCCAC	1375	Db	150728	TTGATTTGGTATTTAAGGCGAGAGTACTTGGCTTTGAATGACTGTATATGTTATTTTA	150787
Db	149666	GGGCTGTGACTAAGGCGGTGGCTTTGTCACTTGACCTCTTAGCT-----TTCTCCAC	149717	QY	2437	CTGTAGCTTAAATATGTTTAAATAGACTTTGAGTGAATGTTTGTGGCAACCTA--GGG	2494
QY	1376	GTGTTTGTGATCTTCTCACTACTCTGTGTGAGATGAGGCAAGGCTTTCAGAGGCA	1435	Db	150788	TTGAGCTTTAAATATATTTTAAATGATTTTGAATTAATGTTTGTGGCAACCTAGGGGG	150847
Db	149718	ATGCTGTGGATCTTCTTCCACTACTCTGTGTGAGACGGAGGCTTGGCTTCAGAGGGCA	149777	QY	2495	GTTTATGATCAGAAATTTGGGCTGAGTCTTTTGGTTTTTGTAGCTTAGAAGTTTTCAGAGG	2554
QY	1436	AGCTGGCAACACCTCTATGCCAAAAAGCTACACTCTCTTTTAAAGCAACATTT--ACATAG	1493	Db	150848	GTTCATGAGCAAGAAATTTGCTGTGAGTCTCTTTGGTTTTTGTAGTCTAGAAATTTTTCAG	150904
Db	149778	AGCCAGTAACACCTCTACGCCAAGAGCTACAGTCCACTTTAAGCAACATTTAAACACTG	149837	QY	2555	GCAAAATCAAACTAGGCTTTGTTTCTGGTTTCACTTACAGGCTCGGAGCTCGCTTATA	2614
QY	1494	ACACTATTTCTGCTTCTCAGAGTGCAGAGCTCTAGACCCACAGAGAAATCTCTCAGGTT	1553	Db	150905	-----ACTAGCTTTGCTTCCAGCTCATCTCACAGGCAAGGGGCTGCTCTTATA	150954
Db	149838	ACATTTATGTCTGCTTCTCAGAGTGCAGAGCTCTAGACCCACAGAGGCTCTCTCTG--	149895	QY	2615	TATTTCTACATAAGGACTTATTATACATAAGTCTGTATAAATGTCTCTGAAGATGACACCTA	2674
QY	1554	ATGTGAAGATTCACACACAGCCCTCT-----CTTGTGACCTCTGTGAGCAAGC	1604	Db	150955	TATTTCTACA-----CGATATACATCTCTGAGGATGACGCCGA	150992
Db	149896	-TGTGGCAATTCACACATAGCCCCCTTACCGTTTGCACTTGACCTCTGTGAGAGCAAGG	149954	QY	2675	GTGCTTTCATCTGAAGGGTCT-----GTCTGGGCTCGAGAGTTGGTTTCAACAGTTAA	2728


```

Db      150993  GCTGCTTTGCTGAAGAGGCGCCGAGAGTCTCTGGGGCTGGAGAGCTGGTTTCAGCCGATAA 151052
QY      2729  GAACCGGTGTTGCTCTCTCCAG-----ACGACCCAGCGACTTGTGAGGCTCAT 2777
Db      151053  AAACACATGTTGCCCTCCAGAGGACCCAGGTTGATTTCCAGCATCTTATAAGAGTGAC 151112
QY      2778  AACACAGAACTCAGCCCTGGGGCATCTGACGCCCTCTCTGACTTCTAAG-----2829
Db      151113  AACCTGTAGCTCCAGTCCAGGGGATCCAGTGGCCCTCTCTGACCTCCAAGGGCAAGAG 151172
QY      2830  ---GCATTTGGTGCACCTACATACAGGCGGCAAAACATTATATACATGTAAGTAATA 2886
Db      151173  CATGATGTTGGTACCTATGATCATACAGGCGCAAGTCAITTCGTAATGTAAGT----151228
QY      2887  AATGCATAGTTAGTGAGAGCGGTGAAGGAAAGGAGTTTT 2926
Db      151229  ATAGTATAGTTAATAAGCGCGGAGGAAAGGTTTTT 151268

RESULT 7
LOCUS   CQ842051                2852 bp      DNA      linear      PAT 02-AUG-2004
DEFINITION   Sequence 698 from Patent EP1440981.
ACCESSION   CQ842051
VERSION     CQ842051.1  GI:50893838
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS    Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J., Isono,Y., Nagai,K. and Irie,R.
TITLE      Full-length human cdna
JOURNAL    Patent: EP 1440981-A 698 28-JUL-2004;
            Research Association for Biotechnology (JP)
FEATURES
            Location/Qualifiers
            source          1..2852
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"

ORIGIN
Query Match      17.3%; Score 722.2; DB 6; Length 2852;
Best Local Similarity 84.8%; Pred. No. 4.1e-184;
Matches 822; Conservative 0; Mismatches 143; Indels 4; Gaps 1;

QY      10  GCTGCTCGATGCTCCAGAGGGCGCATGGGGCCCGCACTGGTGGGACCACTGGCGGC 69
Db      134  GCTGCTCAATGCCCGGAGTGGCCATGGGGCCCGCCGCACTGGTGGGACCACTGGCGGC 193
QY      70  TGGCAGTTCGGAGGTGATTTGGTGGGAGGACAACTACACTATCGTGCCTGCCATTCGCGA 129
Db      194  TGGTAGTTCGGAGGTGACTGGTGGGAGGACAACTACACCATCGTGCCTGCTATCGCGCA 253
QY      130  GTTCTACAAACAGATCAGCAAGCTGTTGTTTTCATTTTACCTCCCATCTGCATGTCCTT 189
Db      254  GTTCTACAAACAGATCAGCAAGCTGTTTTCATTTTACCTCCCATCTGCATGTCCTT 313
QY      190  GTTCCGCGAGTACGCAAGCTGTTTCAACAGCGGCATCTACTATATGAGGCTCCTAGT 249
Db      314  GTTTCGTCAGTATCAACATGCTTCAACAGTGGCATCTACTTAATCTGGACTCTTTGGT 373
QY      250  TGTAGTGGGATTCGATCTGCTACTTCTCCATGCCAAGCTGAGTTTCCTGGGTCAAGTCT 309
Db      374  TGTAGTGGGAATGGATCCGTTCTACTTCCATGCCAAGCTGAGTTTCCTGGGTCAAGTCT 433
QY      310  TGATGAATCTGCCATCTCTGGGTTCTGATGTGCTTTGGCCATGTTGTTTCCAGAG 369
Db      434  TGATGAATCTGCCATCTCTGGGTTCTGATGTGCTTTGGCCATGTTGTTTCCAGAG 493
QY      370  GTATTTACCAAGATCTTTCCGATGACAGGGGAGTTCAAGGAGTGGTGTGTCCT 429

```

```

Db      494  GTATCTACCAAGATCTTTCCGAATGACCGGGGTAGTTTCAAGTGTGGTCAAGTCTCCT 553
QY      430  GTCTCAATTACAAAGTGTGGCTTTATCAAGCCCGCATCAACAATATTTCCTGAT 489
Db      554  GTCTCGGTAGCAGGTGCCCGCATTTGTCAAGCGCTGCATCAACAATCTCTCTGAT 613
QY      490  GATTCTGGGACTTCCATGCACATGCGCTGCTGTGTCAGAGCTGAAGAGGTGTGACAATGT 549
Db      614  GACCTTGGGAGTTCCTTTCGACTGCACTGCTCATTCGAGAGCTAAAGAGGTGTGACAAT 673
QY      550  GCGTGTGTTTAAAGCTGGGCTCTTCTCTGGCCCTCTGGTGGACTTGGCTCTCTTCTGCTG 609
Db      674  GCGTGTGTTTAAAGCTGGGCTCTTCTCTGGGCTCTGCTGGTGGACCTGCGCCCTGTTCTGCTG 733
QY      610  GATCAGCGAACCAAGCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 669
Db      734  GATCAGTGAACCGAGCTTCTGCGAGCTGCTGTCACTCTTCAACTTCCCTTACCTGACCTG 793
QY      670  TGTGTGGCATATTCTCATCTGCTTCTGTTCTGTTACCTGGGCTGTGTGCTTCCCTTACTTT 729
Db      794  CATGTGGCACAATCTCATCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTT 853
QY      730  TGATGCTGCTCAGAGATACCTGAGCAAGTCCAGTCAATCAGATTCTGCGCCCAAGCAGAA 789
Db      854  TGATGCTGCTCAGAGATTCCTGAGCAAGGCTGTCATCAAGTCTCTGCGCCCAATGAGAA 913
QY      790  ATGGGGCTTTTATGTTGCTGCTTATGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 849
Db      914  ATGGGGCTTCTCATTTGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 973
QY      850  CAGATCAGTGTGATGCGCAAGGAGTGCACAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 909
Db      974  CAGATCAGTGTGATGCGCAAGTGTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1029
QY      910  GCTGGGCTTCTGTTGTAGCAAGATGGCTGAGGGGTTGAGGAATTTGGTGTGGTGGG 969
Db      1030  AGTGGGCTTCTTCTGCTAGGAGACAGCAAGGAGTTCGAATAGTTGGGTGTGGGCTA 1089
QY      970  TGTTTAAAA 978
Db      1090  TCTTTTCAA 1098

RESULT 8
LOCUS   AK123581                2852 bp      mRNA      linear      PRI 09-SEP-2003
DEFINITION   Homo sapiens cdna FLJ41587 fis, clone CTONG2202638.
ACCESSION   AK123581
VERSION     AK123581.1  GI:34529163
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS    Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,K., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE      NEDO human cdna sequencing project
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 2852)
AUTHORS    Isogai,T. and Yamamoto,J.
TITLE      Direct Submission
JOURNAL    Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
            Kazuma-Kamatairi, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
            NEDO human cdna sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cdna full insert sequencing:
            Research Association for Biotechnology (RAB); cdna library

```

construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Source

Location/Qualifiers
 1..2852
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CTONG2020638"
 /tissue_type="tongue, tumor tissue"
 /clone_lib="CTONG2"
 /note="cloning vector: pME18SFL3"

ORIGIN

Query Match 17.3%; Score 722.2; DB 9; Length 2852;
 Best Local Similarity 84.8%; Pred. NO. 4.1e-184;
 Matches 822; Conservative 0; Mismatches 143; Indels 4; Gaps 1;

QY 10 GCTGCTCCGATGCTCCAGAGCGCCATGGGCGCCGCTACTGGTGGGACACCTGCGGC 69
 DB 134 GCTGCTCCGATGCTCCAGAGCGCCATGGGCGCCGCTACTGGTGGGACACCTGCGGC 193

QY 70 TGGCAGTTCGAGGTGGATTGGTGGGAGCACTACACTATCGTGCCTGCCATTGCCGA 129
 DB 194 TGGTAGTCGAGGTGGATTGGTGGGAGCACTACACTATCGTGCCTGCCATTGCCGA 253

QY 130 GTTCTACACAGCATCAGCAAGCTGTTGTTTCTATTTTACCTCCCATCTGCATGCTT 189
 DB 254 GTTCTACACAGCATCAGCAAGCTGTTGTTTCTATTTTACCTCCCATCTGCATGCTT 313

QY 190 GTTCCGCGAGTACGCAAGCTGCTTCAACAGCGGCATCTACTTAATATGACGCTCTAGT 249
 DB 314 GTTCTCAGTATGCAACAGCTTCAACAGTGGCATCTACTTAATATGACGCTCTAGT 373

QY 250 TGTAGTGGGATTTGATCTGCTACTTCCATCAACGCTGAGTTTCCGGGTGAGATGCT 309
 DB 374 TGTAGTGGGATTTGATCTGCTACTTCCATCAACGCTGAGTTTCCGGGTGAGATGCT 433

QY 310 TGATCAACTTGCATCTGTGGGTTCTGATGTTGCTTTGGCCATGTTGCTTCCAGGAG 369
 DB 434 TGATCAACTTGCATCTGTGGGTTCTGATGTTGCTTTGGCCATGTTGCTTCCAGGAG 493

QY 370 GTATTTACCAAGATCTTTGCGAATGACAGGGGAGGTTCAAGCAGTGGTGTGTCCT 429
 DB 494 GTATTTACCAAGATCTTTGCGAATGACAGGGGAGGTTCAAGCAGTGGTGTGTCCT 553

QY 430 GTCTGCAATTAACAGCTGTTGGCGTTTATCAAGCCCGCCATCAACAATATTTCCCTGAT 489
 DB 554 GTCTGCGGTTACGAGCTGCGCGCATTTGTCAAGCCTGCCATCAACAATATTTCCCTGAT 613

QY 490 GATTCGGGATTTCCATGCTGCTGCTGTTGTTGAGAGCTGGAAGGTTGCAATGT 549
 DB 614 GATTCGGGATTTCCATGCTGCTGCTGTTGTTGAGAGCTGGAAGGTTGCAATGT 673

QY 550 GCGTGTGTTTAAAGCTGGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
 DB 674 GCGTGTGTTTAAAGCTGGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 733

QY 610 GATCAGCGACCAAGCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 669
 DB 734 GATCAGCGACCAAGCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 793

QY 670 TGTGTGCATATTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
 DB 794 CATGTGGGACATCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853

QY 730 TGATGCTGCTTCAGAGATACCTGAGCAAGTCCAGTCAATCAGATTCTGGGCCAGAGAA 789
 DB 854 TGATGCTGCTTCAGAGATACCTGAGCAAGGCTCTGATCAAGTTCTGGGCCAATGAGAA 913

QY 790 ATGGGCTTTTATTTGGTGTCCCTTATGTTGTCCTTCTGTGTGCCCAAGAGTCCGAGT 849

Db 914 ATGGGCTTTTATTTGGTGTCCCTTATGTTGTCCTTCTGTGTGCCCAAGAAATCATCAGT 973
 QY 850 CAAGATCAGTGTGATGCAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
 Db 974 CAAGATCAGTGTGATGCAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1029
 QY 910 GTTGGGCTTCTGTTTGTAGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
 Db 1030 AGTGGGCTTCTCTTGTAGCAAGACAGCAAGGAGTGTGCTGCTGCTGCTGCTGCTG 1089
 QY 970 TGTGTTAAA 978
 Db 1090 TCTTTCAA 1098

RESULT 9
 AC017081/c
 LOCUS 149462 bp DNA linear PRI 09-JAN-2002
 DEFINITION Homo sapiens BAC clone RP11-470J24 from 2, complete sequence.
 AC017081
 ACCESSION
 VERSION
 KEYWORDS HTG.
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 149462)
 Sultston, J.E. and Waterston, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 99633792
 9847074

REFERENCE
 2 (bases 1 to 149462)
 Nguyen, C., Doebber, A. and Kozlowicz, A.
 The sequence of Homo sapiens BAC clone RP11-470J24
 Unpublished (2001)

REFERENCE
 3 (bases 1 to 149462)
 Waterston, R.H.
 Direct Submission
 Submitted (09-DEC-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE
 4 (bases 1 to 149462)
 Waterston, R.
 Direct Submission
 Submitted (09-JAN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jan 9, 2002 this sequence version replaced gi:14185368.

COMMENT

 Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu

 Summary Statistics

 Center project name: H_NH0470J24

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.

Query Match		16.9%; Score 705; DB 9; Length 149462;
Best Local Similarity		84.4%; Pred. No. 2.7e-179;
Matches		818; Conservative 0; Mismatches 145; Indels 6; Gaps 2;
QY	10	GCTGCTCCGATGTCAGAGCGCCATGGGCGCCGCACTGTGTGGACCACTTGGGGC 69
DB	136946	GCTGCTCCGATGTCAGAGCGCCATGGGCGCCGCACTGTGTGGACCACTTGGGGC 136887
QY	70	TGCGAGTTCGGAGTGGATGGCGAGCACTACACTATGTCGCTGCGCATTTGCCGA 129
DB	136886	TGCGAGTTCGGAGTGGATGGCGAGCACTACACTATGTCGCTGCGCATTTGCCGA 136827
QY	130	GTCTCAACACAGATCAGCAACGCTTGTGTTTTCAITTTTACCTCCCACTGCAATGCTT 189
DB	136826	GTCTCAACACATGATCAGCAATGCTTATTTTCAITTTTACCGCCCATGCAATGCTT 136767
QY	190	GTTCGCGCAGTAGCGAAGCTGCTTCAACAGCGGCATCTACTTAATGAGCGCTCTAGT 249
DB	136766	GTTCGCGCAGTAGCGAAGCTGCTTCAACAGCGGCATCTACTTAATGAGCGCTCTAGT 136709
QY	250	TGTAGTGGGATTTGGATCTGTCTACTTCCATGCAAGCTGAGTTTCTCTGGTCCAGATGCT 309
DB	136708	TGTAGTGGGATTTGGATCTGTCTACTTCCATGCAAGCTGAGTTTCTCTGGTCCAGATGCT 136649
QY	310	TGATGAATCTGCAATCTGTGGGTTCTGATGTGCTTTTGGCCAGTGTGTTCCCAAGGAG 369
DB	136648	TGATGAATCTGCAATCTGTGGGTTCTGATGTGCTTTTGGCCAGTGTGTTCCCAAGGAG 136589
QY	370	GTATTTACCAAGATCTTTGGGAATGACAGGGCGAGGTTCAAGGCGAGTGTGTGTCCT 429
DB	136588	GTATTTACCAAGATCTTTGGGAATGACAGGGCGAGGTTCAAGGCGAGTGTGTGTCCT 136529
QY	430	GTCTGCAATTTACAGCTGTGTGGGTTTATCAAGCCCGCATCAACAATATTTCCCTGAT 489
DB	136528	GTCTGCAATTTACAGCTGTGTGGGTTTATCAAGCCCGCATCAACAATATTTCCCTGAT 136469
QY	490	GATTTGGGACTTCCAGTACGCTGCTGTTGTGAGAGCTGGAAGAGTGTGACAAATGT 549
DB	136468	GATTTGGGACTTCCAGTACGCTGCTGTTGTGAGAGCTGGAAGAGTGTGACAAATGT 136409
QY	550	GGGTGTTTAACTGGGCTCTTCTCTGGGCTGTGGTCTGGTCTGCTCTCTCTGCTG 609
DB	136408	GGGTGTTTAACTGGGCTCTTCTCTGGGCTGTGGTCTGGTCTGCTCTCTCTGCTG 136349
QY	610	GATCAGCGACCAAGCTCTCTGTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 669
DB	136348	GATCAGTGAACGAGCTTCTGTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136289
QY	670	TGTGCGCATATTTCTCATCTGCTTGTCTGTTACCTGGGCTGTGTGCTTGGCCTACTT 729
DB	136288	CATGTGGCACAATCTCTCATCTGCTTGTCTGTTACCTGGGCTGTGTGCTTGGCCTACTT 136229
QY	730	TGATGCTGCTCAGAGATACCTGAGCAAGGTCAGTCAATGATTTGCGCCAGCGAGAA 789
DB	136228	TGATGCTGCTCAGAGATACCTGAGCAAGGTCAGTCAATGATTTGCGCCAGCGAGAA 136169
QY	790	ATGGGCTTTTATTTGTTGCTCCCTTATGTTGCTCCCTTCTGTGTCGCCAAGAGTCGCCAGT 849
DB	136168	ATGGGCTTTTATTTGTTGCTCCCTTATGTTGCTCCCTTCTGTGTCGCCAAGAGATCATAGT 136109
QY	850	CAAGATCAGTGAATGGCAAGGAGTGA CAGGCTTCTCTACTTACTTCTTATTCGATGCGC 909
DB	136108	CAAGATCAGTGAATGGCAAGGAGTGA CAGGCTTCTCTACTTACTTCTTATTCGATGCGC 136053
QY	910	GCTGGGCTTCTGTTGCTAGCAAGATGCTGAGGGGGTGTGAGGATTTGGTGTGGTGGG 969
DB	136052	AGTGGGCTTCTGTTGCTAGCAAGATGCTGAGGGGGTGTGAGGATTTGGTGTGGTGGG 135993
QY	970	TGTTTAAAA 978
DB	135992	TCTTTTCAA 135984

RESULT 10			
AX206799			
LOCUS			
Sequence 6 from Patent WO0155408.			
AX206799			
VERSION			
AX206799.1			
GI:15394643			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1			
Hofmann, K. and Conradt, M.			
Ceramidase			
Patent: WO 0155408-A 6 02-AUG-2001;			
Memorec Medical Molecular Research Cologne Stoffel GmbH (DE)			
Location/Qualifiers			
1..822			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
725; Conservative 0; Mismatches 97; Indels 3; Gaps 1;			
QY	35	ATGGGGCCCCGACCTGTGTGGGACCACTGCGGCTGGCAGTTCGGAGGTGGATTCGTGC	94
Db	1	ATGGGGCCCCGACCTGTGTGGGACCACTGCGGCTGGCAGTTCGGAGGTGGATTCGTGC	60
QY	95	GAGGACAACTACACTATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	154
Db	61	GAGGACAACTACACTATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	120
QY	155	TTGTTTTTCATTTTACCTCCCATCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	214
Db	121	TTATTTTTTCATTTTACCTCCCATCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	180
QY	215	AACAGCGGATCTACTTAATATGAGGCTCTCTAGTTGTAGTGGGAGTTGGAATCTGCTAC	274
Db	181	AACAGCGGATCTACTTAATCTG---GCTCTTGGTTGTAGCGGGAAATTGGATCCGCTCTAC	237
QY	275	TTCCATGCAACGCTGAGTTTCTGGGTCAAGTCTGCTGATGAACTTGCATCTCTTGGGTT	334
Db	238	TTCCATGCAACCTTAGTTTCTGGGTCAAGTCTGCTGATGAACTTGCAGTCTCTTGGGTT	297
QY	335	CTGATGTGTGCTTTGGCCATGTGTTTCCAGGAGGTATTTACCAAGATCTTTCCGGAAT	394
Db	298	CTGATGTGTGCTTTGGGTCTGATGTTCCCGAAGGTATCTACCAAGATCTTTCCGGAAT	357
QY	395	GACAGGGGAGGTTCAAGGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	454
Db	358	GACACGGGTAGGTTCAAGGT	417
QY	455	TTTATCAAGCCCGCCATCAACAATTTTCCCTGATGATTCTGGGACTTCCATGCACTGG	514
Db	418	TTTGTCAAGCCTTGCCATCAACAATCTCTGATGACCTTGGGAGTTCTTTGCGCTGCA	477
QY	515	CTGCTTGTGTGAGAGCTGAAGAGGTGTGACAAATGTGCGTGTGTTTAAAGCTGGGCTCTTC	574
Db	478	CTGCTCATCACAGAGCTAAAGAGGTGTGACAAATGCGTGTGTTTAAAGCTGGGCTCTTC	537
QY	575	TCTGGGCTCTGTGTGACTTGGCTCTCTCTCTGCTGGATCAGGACCAAGCCTTCTGTGAG	634
Db	538	TGGGGCTCTGTGTGACCTTGGCCCTGTCTGCTGGATCAGTGACCGAGCTTCTCTGCGAG	597
QY	635	CTGCT	694
Db	598	CTGCTGTGATCCTTCAACTTCCCTTACTGTCACTGCACTGTGTGGCATATTCTCATCTGC	657
QY	695	GCTTCGTACCTGGGCTGTGTGCTTCCGCTACTTTGATGCTGCTGCTCAGAGATACCTGAG	754

```
Db 658 GGTGCTACTGCTGGCTGTGTATGCTTTGGCTACTTTGATGCTGCTCCCTCAGAGATTCCTGAG 717
QY 755 CAAGGTCCAGTCATCAGATTTCTGGCCAGCAGCAAAATGGCTTTTATTTGTTGCTCCCTTAT 814
Db 718 CAAGGCTGTGTCATCAAAATCTGGCCAGCAGCAAAATGGCTTTTATTTGTTGCTCCCTAT 777
QY 815 GTGTCCCTCTGTGTGTCGCCCAAGAAGTGCCTCAGTCAAGATCAG 859
Db 778 GTGTCCCTCTGTGTGTCGCCCAAGAAGTGCCTCAGTCAAGATCAG 822

RESULT 11
G76816
LOCUS 531 bp DNA linear STS 06-SEP-2002
DEFINITION S208P6606FG12.T0 129S1/SvimJ Mus musculus STS genomic, sequence
tagged site.
ACCESSION G76816
VERSION G76816.1 GI:22727572
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 531)
Wade, C.
Polymorphism Structure in the Mouse
Unpublished (2002)
COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersli@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 531
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGSV3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.
FEATURES
source
Location/Qualifiers
1..531
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S1/SvimJ"
/db_xref="taxon:10090"
/map="4.20-461.84241674-84242112"
/clone_lib="129S1/SvimJ"
<1..531

STS
ORIGIN
Query Match 12.2%; Score 510.8; DB 11; Length 531;
Best Local Similarity 99.1%; Pred. No. 8.4e-127;
Matches 525; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
QY 792 GGGCTTTTATTTGTTGTCCTTATGTGTCCTTCTGTGTGCCCAAGAAGTGCCTCAGTCA 851
Db 1 GGGCTTTTATTTGTTGTCCTTATGTGTCCTTCTGTGTGCCCAAGAAGTGCCTCAGTCA 60
QY 852 AGATCAGTGATGCGAAGGAGTGACAGCTTCTCTACTTACTTCTATTCAGTGCCTG 911
Db 61 AGATCAGTGATGCGAAGGAGTGACAGCTTCTCTACTTACTTCTATTCAGTGCCTG 120
QY 912 TGGGCTTCGTTGCTAGCAAGATGGCTGAGGGGGTTGAGGAATTGTGTGCTGGGTG 971
Db 121 TGGGCTTCGTTGCTAGCAAGATGGCTGAGGGGGTTGAGGAATTGTGTGCTGGGTG 180
```

```
QY 972 TTTAAATTTCTGCTCCTTTGTGATCTAACTGACCACTGTCCTGACCTCCCTAGGTTAA 1031
Db 181 TTTAAATTTCTGCTCCTTTGTGATCTAACTGACCACTGTCCTGACCTCCCTAGGTTAA 240
QY 1032 GTAGAGGGCTCAGGGACATAAGGTGCTTCTCCTCAGTATCCTTTCCAGGCATACGGGCTTG 1091
Db 241 GTAGAGGGCTCAGGGACATAAGGTGCTTCTCCTCAGTATCCTTTCCAGGCATACGGGCTTG 300
QY 1092 CTGGGTTATGTCCATAATGACATCAACAGAGTAGTCTTTTGGGAGACCTTAGGGCAACCCA 1151
Db 301 CTGGGTTATGTCCATAATGACATCAACAGAGTAGTCTTTTGGGAGACCTTAGGGCAACCCA 360
QY 1152 AAGTTTCTTGTGAGGAGGTAGCTTGTCTGT-----TTCCACCATTTCCAGACTCTCAACCC 1208
Db 361 AAGTTTCTTGTGAGGAGGTAGCTTGTCTGTAGTATCCACCATTTCCAGACTCTCAACCC 420
QY 1209 CCTGATAAACAAGACCTTCTGATTGTTGATGAAGGTTCCAGAACTTTTCATTTTGGCG 1268
Db 421 CCTGATAAACAAGACCTTCTGATTGTTGATGAAGGTTCCAGAACTTTTCATTTTGGCG 480
QY 1269 GGAGAAACTGCTCTCAACAAACCAAGTGGGCAAAACACGTGTGGGGG 1318
Db 481 GGAGAAACTGCTCTCAACAAACCAAGTGGGCAAAACACGTGTGGGGG 530

RESULT 12
AX400443/c
LOCUS AX400443 595 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 119 from Patent WO0210453.
ACCESSION AX400443
VERSION AX400443.1 GI:21336623
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
Elashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 119 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
source
Location/Qualifiers
1..595
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/notes="EMBL/GenBank Accession No. AA848826"

ORIGIN
Query Match 9.6%; Score 400.4; DB 6; Length 595;
Best Local Similarity 87.9%; Pred. No. 7.8e-97;
Matches 531; Conservative 0; Mismatches 56; Indels 17; Gaps 8;
```


Search completed: November 21, 2004, 03:07:23
Job time : 16954.5 secs

Db 62 GGATCAGTGACCGAGCTTTCTGCGAGCTGCTGTATCATCTTCAACTTCCCTACCTGCCTACT 121
Qy 669 GTGTGTGGCATATTCTCATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 728
Db 122 GCATGTGGCACATCTCATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 181
Qy 729 TTGATGCTGCTCAGAGTACCTCAGCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 788
Db 182 TTGATGCTGCTCAGAGTACCTCAGCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 241
Qy 789 AATGGGCTTTTATTGGTGTCCCTTATGTGTCCCTTCTGTGTCCCTTCTGTGTCCCTTCTGTGT 848
Db 242 AATGGGCTTCAATGGTGTCCCTTATGTGTCCCTTCTGTGTCCCTTCTGTGTCCCTTCTGTGT 301
Qy 849 TCAAGATCAGGTGA 862
Db 302 TCAAGACCAAGTGA 315

RESULT 15
CQ736906
LOCUS CQ736906 315 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 22840 from Patent WO02068579.
ACCESSION CQ736906
VERSION CQ736906.1 GI:42333767
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 22840 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1. .315
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 6.3%; Score 261.2; DB 6; Length 315;
Best Local Similarity 89.5%; Pred. No. 4.5e-59;
Matches 281; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Qy 549 TGGTGTGTTTAAAGCTGGGCTCTTCTCTGCGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCT 608
Db 2 TGGTGTGTTTAAAGCTGGGCTCTTCTCTGCGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCT 61
Qy 609 GGATCAGCGACCAAGCCTTCTGTGAGCTGCTCTCTCTCTTCTCACTTCCCTTACCTGCCTACT 668
Db 62 GGATCAGTGACCGAGCTTCTGCGAGCTGCTGTATCATCTTCAACTTCCCTACCTGCCTACT 121
Qy 669 GTGTGTGGCATATTCTCATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 728
Db 122 GCATGTGGCACATCTCATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 181
Qy 729 TTGATGCTGCTCAGAGTACCTCAGCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 788
Db 182 TTGATGCTGCTCAGAGTACCTCAGCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 241
Qy 789 AATGGGCTTTTATTGGTGTCCCTTATGTGTCCCTTCTGTGTCCCTTCTGTGTCCCTTCTGTGT 848
Db 242 AATGGGCTTCAATGGTGTCCCTTATGTGTCCCTTCTGTGTCCCTTCTGTGTCCCTTCTGTGT 301
Qy 849 TCAAGATCAGGTGA 862
Db 302 TCAAGACCAAGTGA 315

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:48:16 ; Search time 1770.81 Seconds
(without alignments)
12376.463 Million cell updates/sec

Title: US-10-017-410-1
Perfect score: 4175
Sequence: 1 ggcacagggctgcgcgat.....aaaaaaaaaaaaaaaaaac 4175

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4175	100.0	4175	6	ABSS54148
2	723.8	17.3	4202	8	ACC50961 Human bla
3	723.8	17.3	4212	11	ADN38833 Cancer/an
4	684.8	15.9	828	6	ABSS54149
5	653.8	15.7	822	4	AHA48641 Human cer
6	424.2	10.2	1623	10	ADI21948
7	400.4	9.6	595	6	ABK62212 Rat seque
8	400.4	9.6	595	10	ADB55226
9	323.6	7.8	1527	10	ADL06640
10	323.6	7.8	1527	10	ADL06642 Human 3T3
11	294.8	7.1	487	10	ADI21468
12	176.8	4.2	1215	4	AAK51566
13	176.8	4.2	1631	4	AAK52550 Human pol
14	174.8	4.2	792	4	AHA48639 Human cer
15	160.6	3.8	537	4	ABL24043
16	125.8	3.0	1473	4	AAH33545 Human col
17	108	2.6	2143	5	ABV24475 Human pro
18	108	2.6	2143	5	ABV24620 Human pro
19	108	2.6	2143	5	ABV24928 Human pro
20	92.6	2.2	4758	4	ABL24042 Drosophil
21	77.8	1.9	4071	4	ABL30432 Drosophil

c	22	70.6	1.7	692	5	ABV28857	Abv28857 Human pro
c	23	70.6	1.7	692	5	ABV23021	Abv23021 Human pro
c	24	61.4	1.5	110000	8	ABX16390_3	Continuation (4 of
c	25	59	1.4	2000	8	ADA71938	Ada71938 Rice gene
c	26	58.2	1.4	35425	10	ADC87616	Adc87616 Human GPC
c	27	54.6	1.3	14658	12	ADM80713	Adm80713 Mouse USH
c	28	52.6	1.3	19619	6	ABK47192	Abk47192 Mouse Con
c	29	49.8	1.2	2000	8	ADA71938	Ada71938 Rice gene
c	30	49.4	1.2	7516	6	ABL32061	Ab132061 Human imm
c	31	48.6	1.2	50295	9	ADA02993	Ada02993 Mouse Sep
c	32	48.6	1.2	50295	10	ADB72731	Adb72731 Mouse Sep
c	33	48.6	1.2	50295	10	ADC85473	Adc85473 Mouse Sep
c	34	48.6	1.2	50295	12	ADM74588	Adm74588 Murine ca
c	35	47.6	1.1	52479	9	ADA02795	Ada02795 Mouse Tnf
c	36	47.6	1.1	52479	10	ADB72533	Adb72533 Mouse Tnf
c	37	47.6	1.1	52479	10	ADC85275	Adc85275 Mouse Tnf
c	38	47.6	1.1	52479	12	ADM74390	Adm74390 Murine ca
c	39	47.4	1.1	215980	6	ABL38337	Ab138337 Complemen
c	40	47	1.1	5734	10	ADC85382	Adc85382 Human Ptp
c	41	47	1.1	67191	9	ADA02903	Ada02903 Mouse Mcm
c	42	47	1.1	67191	10	ADB72641	Adb72641 Mouse Mcm
c	43	47	1.1	67191	11	ADL95495	Adl95495 Murine MC
c	44	47	1.1	67191	12	ADM74498	Adm74498 Murine ca
c	45	46.6	1.1	171936	6	ABS56565	AbS56565 Human SUL

ALIGNMENTS

RESULT 1	
ABS54148	
ID	ABS54148 standard; cDNA; 4175 BP.
XX	AC ABS54148;
XX	AC ABS54148;
DT	25-NOV-2002 (first entry)
XX	DE Mouse cDNA encoding a liver tumour marker protein, CRG-L1.
XX	XX Mouse; ss; gene; liver cancer; liver tumour; CRG-L1;
KW	hepatocellular cancer.
KW	hepatocellular cancer.
OS	Mus musculus.
XX	XX
FH	Key Location/Qualifiers
FT	CDS 35..862
FT	/*tag= a
FT	/product= "Liver tumour marker protein, CRG-L1"
XX	US2002115094-A1.
XX	22-AUG-2002.
PD	14-DEC-2001; 2001US-00017410.
XX	14-DEC-2001; 2001US-00017410.
PR	14-DEC-2000; 2000US-0255674P.
XX	(FARN/) FARNHAM P J.
PA	(GRAV/) GRAVEEL C R.
XX	Farnham PJ, Graveel CR;
FI	WPI; 2002-706409/76.
XX	P-PSDB; ABG32880.
DR	Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
XX	liver cancer, is differentially expressed in liver tumours relative to
PT	normal liver tissues.
PT	Claim 2; Page 4-7; 11pp; English.
PS	The invention relates to a polypeptide designated as CRG-L1, which is
XX	differentially expressed in liver tumours relative to normal expression
CC	

CC in normal liver tissues, designated CRG-L1. Also included are the
 CC encoding polynucleotides (in the case of the human sequence, mapping to
 CC chromosome 9p), expression constructs, host cells, anti-CRG-L1
 CC antibodies, identifying modulators of CRG-L1, and the use of the CRG-L1
 CC sequence in the diagnosis of hepatocellular cancer in tumour cells from a
 CC liver of a human or non-human animal. The CRG-L1 protein and
 CC polynucleotide are useful as diagnostic markers for a liver cancer in
 CC humans and non-human animals, and as a system for assessing putative
 CC therapeutic agents. The present sequence encodes mouse CRG-L1
 XX

Sequence 4175 BP; 962 A; 971 C; 1012 G; 1230 T; 0 U; 0 Other;

Query Match 100.0%; Score 4175; DB 6; Length 4175;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCAGAGGCTGCTCCGATGCTCCAGAGCGGCATGGCGCCCGCACTGGTGGGACCA	60
DB	1	GGCAGAGGCTGCTCCGATGCTCCAGAGCGGCATGGCGCCCGCACTGGTGGGACCA	60
QY	61	CTTGCGGGCTGGCAGTTCGGAGTGGATGGTGGGAGGACCACTACACTATCGTCTGC	120
DB	61	CTTGCGGGCTGGCAGTTCGGAGTGGATGGTGGGAGGACCACTACACTATCGTCTGC	120
QY	121	CATTGCGAGTCTCAACACAGATCAGCAAGCTCTTTGTTTTCATTTTACCTCCCATCTG	180
DB	121	CATTGCGAGTCTCAACACAGATCAGCAAGCTCTTTGTTTTCATTTTACCTCCCATCTG	180
QY	181	CATTGCTTTGTTCCGCGAGTACGCAAGCTCTTCAACAGCGGCATCTACTTAATGGAC	240
DB	181	CATTGCTTTGTTCCGCGAGTACGCAAGCTCTTCAACAGCGGCATCTACTTAATGGAC	240
QY	241	GCTCTAGTGTAGTGGGATGGATCTGCTACTTCCATGCAAGCTGAGTTCTCTGG	300
DB	241	GCTCTAGTGTAGTGGGATGGATCTGCTACTTCCATGCAAGCTGAGTTCTCTGG	300
QY	301	TCAGATGCTGTGAATCTGCCATCTCTGGGTTCTGATGTGCTTGGCCATGGTT	360
DB	301	TCAGATGCTGTGAATCTGCCATCTCTGGGTTCTGATGTGCTTGGCCATGGTT	360
QY	361	TCCAGAGGATATTTACAAAGATCTTTCCGAATGACAGGGCGAGTTCAAGSCAGTGT	420
DB	361	TCCAGAGGATATTTACAAAGATCTTTCCGAATGACAGGGCGAGTTCAAGSCAGTGT	420
QY	421	GTGTGCTGTGCAATTTACAAAGCTGTGGGTTTATCAAGCCGCCATCAACAATAT	480
DB	421	GTGTGCTGTGCAATTTACAAAGCTGTGGGTTTATCAAGCCGCCATCAACAATAT	480
QY	481	TTCCCTGATGATTTCTGGGACTTCCATGCTGGCTGCTTGTGAGAGCTGAAGAGTG	540
DB	481	TTCCCTGATGATTTCTGGGACTTCCATGCTGGCTGCTTGTGAGAGCTGAAGAGTG	540
QY	541	TGACAAATGTCGCTGTTTAAAGCTGGGCTCTTCTCTGGCTCTGCTGGACTCTGGCTCT	600
DB	541	TGACAAATGTCGCTGTTTAAAGCTGGGCTCTTCTCTGGCTCTGCTGGACTCTGGCTCT	600
QY	601	CTTCTGCTGGATCAGGACCAAGCTTCTGAGCTGCTCTCTCTCTCTCTCTCTCTCT	660
DB	601	CTTCTGCTGGATCAGGACCAAGCTTCTGAGCTGCTCTCTCTCTCTCTCTCTCTCT	660
QY	661	CCTGCACTGTGTGGGATATTTCTATCTGCTTGTCTGCTGCTGCTGCTGCTGCTT	720
DB	661	CCTGCACTGTGTGGGATATTTCTATCTGCTTGTCTGCTGCTGCTGCTGCTGCTT	720
QY	721	CGCTACTTTGATGCTGCTCAGAGATACCTGAGCAAGGTCCAGTCAATCAGATTCGGCC	780
DB	721	CGCTACTTTGATGCTGCTCAGAGATACCTGAGCAAGGTCCAGTCAATCAGATTCGGCC	780
QY	781	CAGCGAGAAATGGGCTTTTATTTGGTGTCCCTTATGCTCTCTGCTGCTGCTGCTGCT	840
DB	781	CAGCGAGAAATGGGCTTTTATTTGGTGTCCCTTATGCTCTCTGCTGCTGCTGCTGCT	840
QY	841	GTGCGCAGTCAAGATCACGATGATGGGCAAGGAGTGCACGAGCTTCTCTACTTCTATT	900

DB	841	GTGCGCAGTCAAGATCACGATGATGGCAAGCAGTGCACGACTTCTCTACTTCTATT	900
QY	901	CGAGTGGCGCTGGGCTTCGTTTGTAGCAAGATGGCTGAGGGGTGTAGGAATTTGGTG	960
DB	901	CGAGTGGCGCTGGGCTTCGTTTGTAGCAAGATGGCTGAGGGGTGTAGGAATTTGGTG	960
QY	961	TGTTGTGGGTGTTTAAATTTCTGCTCTTTGTGATCTAACTGGACCACTGTGCTGACCT	1020
DB	961	TGTTGTGGGTGTTTAAATTTCTGCTCTTTGTGATCTAACTGGACCACTGTGCTGACCT	1020
QY	1021	CCTAGGTTAAGTAGAGGCTCAGGGACATAAGTGTCTTCTCAGTATCTTTCAGGC	1080
DB	1021	CCTAGGTTAAGTAGAGGCTCAGGGACATAAGTGTCTTCTCAGTATCTTTCAGGC	1080
QY	1081	ATACGGGCTTGTGGGTATGTCATTAATGACATCAACAGAGTAGTTCTTTGGGAGACCT	1140
DB	1081	ATACGGGCTTGTGGGTATGTCATTAATGACATCAACAGAGTAGTTCTTTGGGAGACCT	1140
QY	1141	AGGGCAACCCAAAGTTTCTGCTGGAGAGGTAGTGTCTTCCACCAATTTCCAGACT	1200
DB	1141	AGGGCAACCCAAAGTTTCTGCTGGAGAGGTAGTGTCTTCCACCAATTTCCAGACT	1200
QY	1201	CTCAACCCCTGATATAAACAAGACCTTCTGATTTGGTGTGATAAGAGTTCCAGAACTTTCA	1260
DB	1201	CTCAACCCCTGATATAAACAAGACCTTCTGATTTGGTGTGATAAGAGTTCCAGAACTTTCA	1260
QY	1261	TTTTGCGGGAGAACTGCTTCAACAAACCAAGTGGGCAACCAACGTTGGGGGTG	1320
DB	1261	TTTTGCGGGAGAACTGCTTCAACAAACCAAGTGGGCAACCAACGTTGGGGGTG	1320
QY	1321	TGACTAAGAGGGTGGCTTGTCACTCGCACCTCTTAGCTTCCAGGTTCTCCACGTT	1380
DB	1321	TGACTAAGAGGGTGGCTTGTCACTCGCACCTCTTAGCTTCCAGGTTCTCCACGTT	1380
QY	1381	TGTGATCTTCCACTACCTCTGCTGAGAGTGGAGCCACCGCTTCAAGGGCAAGCTG	1440
DB	1381	TGTGATCTTCCACTACCTCTGCTGAGAGTGGAGCCACCGCTTCAAGGGCAAGCTG	1440
QY	1441	GCAACACCTCTATGCAAAAAGCTACACCTCTTTAAGCACACATTACATGACACTAT	1500
DB	1441	GCAACACCTCTATGCAAAAAGCTACACCTCTTTAAGCACACATTACATGACACTAT	1500
QY	1501	TTCTGCTCTTCCAGAGTGCAGAGCTCAGACCCACAGAGAACTCTTCAAGTTATCTGA	1560
DB	1501	TTCTGCTCTTCCAGAGTGCAGAGCTCAGACCCACAGAGAACTCTTCAAGTTATCTGA	1560
QY	1561	GATTTCCACACACAGCCCTCTTGTGACCTCTGTGAGAGCAAGGCTCGTTGTAATCGCG	1620
DB	1561	GATTTCCACACACAGCCCTCTTGTGACCTCTGTGAGAGCAAGGCTCGTTGTAATCGCG	1620
QY	1621	GAGGAGCTCTGGAGGCTTGAAGTGGAGCGCTGTGAGAGTGGAGAGTTTGGAAAGCTTTT	1680
DB	1621	GAGGAGCTCTGGAGGCTTGAAGTGGAGCGCTGTGAGAGTGGAGAGTTTGGAAAGCTTTT	1680
QY	1681	TTGGAGCTTTTGGTGTGCTGTTTACCTCAAGATTTCTGACCCGCTTCTCTGCTGTTT	1740
DB	1681	TTGGAGCTTTTGGTGTGCTGTTTACCTCAAGATTTCTGACCCGCTTCTCTGCTGTTT	1740
QY	1741	AGGGGATATGTCCTTCTGAGTGAAGTACCTCCCTAGGATGTGGGCTCCGGCTTTTGT	1800
DB	1741	AGGGGATATGTCCTTCTGAGTGAAGTACCTCCCTAGGATGTGGGCTCCGGCTTTTGT	1800
QY	1801	TTTCATATACCTGGTGTAACTGGTTTCTCAGAGCACTTTGCTCTTCTGATGCTGGGC	1860
DB	1801	TTTCATATACCTGGTGTAACTGGTTTCTCAGAGCACTTTGCTCTTCTGATGCTGGGC	1860
QY	1861	GGTCACTACCTGATTTGGCTGGGATCCAGGGAGGAAAGGGGAGAGCTCTGACA	1920
DB	1861	GGTCACTACCTGATTTGGCTGGGATCCAGGGAGGAAAGGGGAGAGCTCTGACA	1920
QY	1921	TGATGTCTCTCACCTGCACTCTGCTTCACTGGGCTCTCTTCAAGGGTATGAAGTGG	1980

1971 TGAATGTCCTCACTGATCCTACTGCTCTTCACTGGGCTCTCTTCAGGGTATGAATGG 1980
 1981 GTATGGGGTATCTCAGGGATGTTGTGTAATCTCAGGCACCTTCTGCTTTCTGACATTCCTAT 2040
 1981 GTATGGGGTATCTCAGGGATGTTGTGTAATCTCAGGCACCTTCTGCTTTCTGACATTCCTAT 2040
 2041 GTCACTGTGAGATCTGCACTCAGGGGACTGGAGTTGGAACCCGTTCTGAGTGGTGTCT 2100
 2041 GTCACTGTGAGATCTGCACTCAGGGGACTGGAGTTGGAACCCGTTCTGAGTGGTGTCT 2100
 2101 GAGTCTGAAGGAGTGTAGCTGGGGTCTTGGGACCTTCTAGGATCTCTTACTCTGTTTGA 2160
 2101 GAGTCTGAAGGAGTGTAGCTGGGGTCTTGGGACCTTCTAGGATCTCTTACTCTGTTTGA 2160
 2161 CTTTCAAGGTA CAAGTGGGAATCGA CTTAAAGAGTTTAAATGAATAGACTTCAT 2220
 2161 CTTTCAAGGTA CAAGTGGGAATCGA CTTAAAGAGTTTAAATGAATAGACTTCAT 2220
 2221 TCTGCTTTTGGCTTTGGAGTTTAAAGTAACCTTCTGGGCTATTTCTTGTATTAGT 2280
 2221 TCTGCTTTTGGCTTTGGAGTTTAAAGTAACCTTCTGGGCTATTTCTTGTATTAGT 2280
 2281 CACAAATATTTATGAACATGAAGTGTAAATAGTTGTCTTTTAAATTCATAGCA 2340
 2281 CACAAATATTTATGAACATGAAGTGTAAATAGTTGTCTTTTAAATTCATAGCA 2340
 2341 TTTACCAACTCTCCAGGTAGCAACA CACAGTCATTTTGA TTGGGATTTTCAAGGCAGG 2400
 2341 TTTACCAACTCTCCAGGTAGCAACA CACAGTCATTTTGA TTGGGATTTTCAAGGCAGG 2400
 2401 AATACTTGGTTTGAATGATTATATGTCATTTTACTGTAGCTTAAATATGTTTAAA 2460
 2401 AATACTTGGTTTGAATGATTATATGTCATTTTACTGTAGCTTAAATATGTTTAAA 2460
 2461 TGACTTTGAGTGAATGTTTGTGGCAACTAGGGGTTTATGGATCAGAAATTCGGGCTGAG 2520
 2461 TGACTTTGAGTGAATGTTTGTGGCAACTAGGGGTTTATGGATCAGAAATTCGGGCTGAG 2520
 2521 TCCCTTGGTTTGTAGTCTGAAAGTTTTCAGAGGCAAAATCAAACCTAGCTTGTCTG 2580
 2521 TCCCTTGGTTTGTAGTCTGAAAGTTTTCAGAGGCAAAATCAAACCTAGCTTGTCTG 2580
 2581 GTTCATCTTACAGGCTGGGAGCTGCCCTTATATTTCTACATAGGACTTATATACA 2640
 2581 GTTCATCTTACAGGCTGGGAGCTGCCCTTATATTTCTACATAGGACTTATATACA 2640
 2641 TAACTCTGTATAATGTCCTGAAGATGACACTAGCTGCCCTTCACTCGGAAGGGTGTCT 2700
 2641 TAACTCTGTATAATGTCCTGAAGATGACACTAGCTGCCCTTCACTCGGAAGGGTGTCT 2700
 2701 GGGGCTGGAGTGGTTTCAACAGTTTAAAGACCCGTTGTCTCTCTCCAGACGCCAG 2760
 2701 GGGGCTGGAGTGGTTTCAACAGTTTAAAGACCCGTTGTCTCTCTCCAGACGCCAG 2760
 2761 CGACTTGTAGGCTCATAAACAGCACTCCAGGCTGGGCTATCTGAGCCCTCTCTCTGA 2820
 2761 CGACTTGTAGGCTCATAAACAGCACTCCAGGCTGGGCTATCTGAGCCCTCTCTCTGA 2820
 2821 CTTCTAAGGGCATTTGGTGTGACACGTACATACAGGAGGAGTTTATACATGTAAAC 2880
 2821 CTTCTAAGGGCATTTGGTGTGACACGTACATACAGGAGGAGTTTATACATGTAAAC 2880
 2881 GTAATAATGATTAAGTTAGTGAGCGGCTGAAGAAAGAGTTTATAGATCAAGGTTTA 2940
 2881 GTAATAATGATTAAGTTAGTGAGCGGCTGAAGAAAGAGTTTATAGATCAAGGTTTA 2940
 2941 GTCTGACCTGAGTCCCTTCTTCTAGCCGATGTTCTCAAGTCAAGGAGCAACCAT 3000
 2941 GTCTGACCTGAGTCCCTTCTTCTAGCCGATGTTCTCAAGTCAAGGAGCAACCAT 3000
 3001 TTCTTTTATCAGGAAAGGACTTATTTCCCTAGGGCTCTGCTGACATCCCTAGGAACA 3060
 3001 TTCTTTTATCAGGAAAGGACTTATTTCCCTAGGGCTCTGCTGACATCCCTAGGAACA 3060

3061 GAGATAAAATACCGATGGATGTGAATGAATGCTTGGTGAAGAGCCGAGTACCTGA 3120
 3061 GAGATAAAATACCGATGGATGTGAATGAATGCTTGGTGAAGAGCCGAGTACCTGA 3120
 3121 CTGACCCAGTGGGCACTTCCACAGGAGAGCCGGTAAACCGATGTCACCTCCAGATC 3180
 3121 CTGACCCAGTGGGCACTTCCACAGGAGAGCCGGTAAACCGATGTCACCTCCAGATC 3180
 3181 TTCTCCCTTCTGGTAGAGATGACATCGATGATGTCGTGACCACTGGGCCCCGATG 3240
 3181 TTCTCCCTTCTGGTAGAGATGACATCGATGATGTCGTGACCACTGGGCCCCGATG 3240
 3241 GGTGTGGGCTCTGTTCTTCTTCCGAATCTACTGAGATCTCAGGACAGGAGCCATGA 3300
 3241 GGTGTGGGCTCTGTTCTTCTTCCGAATCTACTGAGATCTCAGGACAGGAGCCATGA 3300
 3301 ATGTTACCAAGTGGTCTGCTGTCAGTGAATTTTACAGTTTGTGACCAATTTTGGTTTT 3360
 3301 ATGTTACCAAGTGGTCTGCTGTCAGTGAATTTTACAGTTTGTGACCAATTTTGGTTTT 3360
 3361 AGGAGAAATTTCTTCTCTCTAGTGGCTGTGATGCCAAGCCAGCCCTTCAGAAAGTTC 3420
 3361 AGGAGAAATTTCTTCTCTCTAGTGGCTGTGATGCCAAGCCAGCCCTTCAGAAAGTTC 3420
 3421 CTTCTGTCTGCTTCCCTTTGTAATGTGATCCACTGGGGAATTTGGTGTCTACCCAGGA 3480
 3421 CTTCTGTCTGCTTCCCTTTGTAATGTGATCCACTGGGGAATTTGGTGTCTACCCAGGA 3480
 3481 AAAGTGCCCTACCCAGTCAAGGCTCAAGGTCACAGTGTGTTGAGGATCAACATGGCTCTGT 3540
 3481 AAAGTGCCCTACCCAGTCAAGGCTCAAGGTCACAGTGTGTTGAGGATCAACATGGCTCTGT 3540
 3541 GAAATAACTGCGCACCCATTTCTATTGGTTGGATTTCTCAGGAGTCTGAATCTTCCCTCAG 3600
 3541 GAAATAACTGCGCACCCATTTCTATTGGTTGGATTTCTCAGGAGTCTGAATCTTCCCTCAG 3600
 3601 AGTCCCTTCTTCCCAACCCCTACCCAGAGCCAAACCTGGGATTTGAACCTTCTCTGACTC 3660
 3601 AGTCCCTTCTTCCCAACCCCTACCCAGAGCCAAACCTGGGATTTGAACCTTCTCTGACTC 3660
 3661 TCTTCTCCCTCAGGCTGACAACTAATGGTCTCTGGGACACCCAGCTAGGCTTCCC 3720
 3661 TCTTCTCCCTCAGGCTGACAACTAATGGTCTCTGGGACACCCAGCTAGGCTTCCC 3720
 3721 CAACCTCTTATCCAGCTGAACCTTGGATTTCTTCCCAACCCAGGCTTGAACCTTCTCTGACTC 3780
 3721 CAACCTCTTATCCAGCTGAACCTTGGATTTCTTCCCAACCCAGGCTTGAACCTTCTCTGACTC 3780
 3781 GGTCCCACTGGCCCAACACATCTTTGGCCAGATTTGGGATTTCAATAGATTTTATAGA 3840
 3781 GGTCCCACTGGCCCAACACATCTTTGGCCAGATTTGGGATTTCAATAGATTTTATAGA 3840
 3841 CATTAATTTCCCAAGACTTTTAAACATGGCTTGTGTTTCCCAACCCAGCTAGGCTTCCC 3900
 3841 CATTAATTTCCCAAGACTTTTAAACATGGCTTGTGTTTCCCAACCCAGCTAGGCTTCCC 3900
 3901 TTTTAAACTATTTTATTAACCAAGAAATTTAAACCAAGCAATAGAGTACTTTTCAGATATA 3960
 3901 TTTTAAACTATTTTATTAACCAAGAAATTTAAACCAAGCAATAGAGTACTTTTCAGATATA 3960
 3961 AACTGTGTTTCACTTTATGTAGAGTGTCTATGTATAGCGGATGTACCCCTGGCTGA 4020
 3961 AACTGTGTTTCACTTTATGTAGAGTGTCTATGTATAGCGGATGTACCCCTGGCTGA 4020
 4021 AGTAATTTAACCAATAGCTCTGGGAGGATTTACAGACCTTTTGCACCTTTATGCTTTTTT 4080
 4021 AGTAATTTAACCAATAGCTCTGGGAGGATTTACAGACCTTTTGCACCTTTATGCTTTTTT 4080
 4081 TGAACCTCTGATAACCATGGTCAATATTTAAAGCCAAATTAACCTGGCATTTTCTGTGAATAAAC 4140
 4081 TGAACCTCTGATAACCATGGTCAATATTTAAAGCCAAATTAACCTGGCATTTTCTGTGAATAAAC 4140

fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
inflammatory disease; autoimmune disease;
retinal neovascularisation syndrome; scarring; uterine fibroid;
detection; diagnosis; prognosis; drug screening; drug targeting;
wound healing; contraception; cytostatic; cardiant; immunomodulatory;
vulnerary; gene therapy; vaccine; gene; ss.

Homo sapiens.

WO2003042661-A2.

22-MAY-2003.

13-NOV-2002; 2002WO-US036810.

13-NOV-2001; 2001US-0350666P.

21-NOV-2001; 2001US-0332464P.

29-NOV-2001; 2001US-0334393P.

03-DEC-2001; 2001US-0335394P.

14-DEC-2001; 2001US-0340376P.

08-JAN-2002; 2002US-0347211P.

10-JAN-2002; 2002US-0347349P.

08-FEB-2002; 2002US-0355250P.

13-FEB-2002; 2002US-0356714P.

20-FEB-2002; 2002US-0359077P.

29-MAR-2002; 2002US-0368809P.

04-APR-2002; 2002US-0370110P.

12-APR-2002; 2002US-0372246P.

05-JUN-2002; 2002US-0386614P.

16-JUL-2002; 2002US-0396839P.

22-JUL-2002; 2002US-0397775P.

22-JUL-2002; 2002US-0397845P.

09-SEP-2002; 2002US-0409450P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Heverzi PA;

Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

WPI; 2003-469649/44.

P-FSDB; ADN38834.

Determining the presence or absence of a pathological cell in a patient,
useful for diagnosing, prognosing or treating cancer, comprises detecting
a nucleic acid in a biological sample.

Claim 8; SEQ ID NO 151; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
whose expression is upregulated or downregulated in specific cancers or
other diseases such as angiogenic or fibrotic disorders, and to methods
of determining the presence or absence of a pathological cell in a
patient by detecting a nucleic acid at least 80% identical to those of
the invention or by detecting a polypeptide of the invention. The
invention also relates to expression vectors and host cells comprising a
nucleic acid of the invention; antibodies which specifically bind a
polypeptide of the invention; use of such antibodies for drug targeting;
and methods of screening for modulators of activity or expression of the
polypeptides and nucleic acids. The nucleic acids, polypeptides,
antibodies and methods are useful for diagnosing, prognosing and treating
cancer and other conditions such as psoriasis, ischaemia, heart disease,
atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
neovascularisation syndromes, scarring and uterine fibroids. They may
also be useful in wound healing and in contraception. The present
sequence represents a nucleic acid sequence of the invention.

Query Match 17.3%; Score 723.8; DB 11; Length 4212;
Best Local Similarity 84.9%; Pred. No. 5.5e-189;
Matches 823; Conservative 0; Mismatches 142; Indels 4; Gaps 1;
QY 10 GCTGCTCCGATGCTCCAGAGCGGCATGGGGCCCGCCGACCTGGTGGACACCTGGCGGC 69

Db	62	GCTGCTCCAAATGCCCGGAGTGGCCATGGCGCCGCACTGGTGAGACAGCTGCAGC	121
QY	70	TGGCAGTTCGGAGGTGATTTGGTGGCAGACAACTACACTATCTGCTGCTCCATTCGCGA	129
Db	122	TGGTAGCTCGAGGTGGACTGGTGGCAGGACAACTACACCATCTGCTGCTATCGCCGA	181
QY	130	GTTCTACACACGATCAGCAACGCTCTTGTCTTTTTCATTTTACCTTCCCATCTGCACTGTCTT	189
Db	182	GTTCTACACACGATCAGCAATGCTTATTTTTCATTTTACCGCCCATCTGCACTGTCTT	241
QY	190	GTTCCGCCAGTACGCAACGCTGCTTCAACAGCGGCACTACTTAATAATGACGCTCTCTAGT	249
Db	242	GTTTCGTAGTATGCAACATGCTTCAACAGTGGCATCTACTTAATCTGACTCTTTTGGT	301
QY	250	TGTAGTGGGGATTGGATCTGCTACTTCCATCGACGCTGAGTTTCTGGGTGAGATGCT	309
Db	302	TGTAGTGGGAATGGATCGCTCTACTTCCATGCAACCCCTTAGTTTCTGGGTGAGATGCT	361
QY	310	TGATGAACCTTGCCATTCTGTGGGTTCTGATGTGTCTTGGGCACTGTGGTTCCTCCAGGAG	369
Db	362	TGATGAACCTTGCACTCTTGGGTTCTGATGTGTCTTGGGCACTGTGGTTCCTCCAGGAG	421
QY	370	GTATTTACAAAGATCTTTCGGAATGACAGGGGCAAGTTCAAGGCACTGTGTGTCTCT	429
Db	422	GTATCTACCAAGATCTTTCGGAATGACCGGGTAGTTTCAAGTGGTGGTCACTGTCTCT	481
QY	430	GTCTCAATTAACAAGTCTTGGCGCTTATCAAGCCCGCATCAACAATATTTCCCTGAT	489
Db	482	GTCTCGGTGACAGCTGCTGGCATTTGTCAAGCCTGCCATCAACAACATCTCTGAT	541
QY	490	GATTCGGGACTTCCATGCACTGCGCTGCTTGTTCAGAGCTGAAGAGGTGCAACATCT	549
Db	542	GACCTCGGAGTTCTTGCATGCACTGCTCATCCAGAGCTAAAGAGGTGCAACAT	601
QY	550	CGGTGTTTAAAGCTGGGCTCTTCTGGGCTCTGGTGGACTCTGGTCTCTTCTGCTG	609
Db	602	CGGTGTTTAAAGCTGGGCTCTTCTGGGCTCTGGTGGACTCTGGGCTCTTCTGCTG	661
QY	610	GATCAGCAGCAAGCTTCTGTGAGCTGCTCTCCCTCTTCACTTCCCTTACCTGCACTG	669
Db	662	GATCAGTACCGAGCTTCTCGAGCTGCTGTCATCTCTTCAACTTCCCTTACCTGCACTG	721
QY	670	TGTGTGCAATATCTCATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	729
Db	722	CATGTGGCAGATCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	781
QY	730	TGATGCTGCTCAGAGATACCTGAGCAAGGTTCAGTCAATCAGATTCCTGGCCCGAGGAA	789
Db	782	TGATGCTGCTCAGAGATTCCTGAGCAAGGCTCTGTCATCAAGTTCTGGCCCAATGAA	841
QY	790	ATGGGCTTTTATGCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	849
Db	842	ATGGGCTTTTATGCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	901
QY	850	CAAGATCAGTGTGCAAGGCTGACCACTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	909
Db	902	CAAGATCAGTGTGCAAGGCTGACCACTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	957
QY	910	GCTGGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	969
Db	958	AGTGGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1017
QY	970	TGTTTAAA 978	
Db	1018	TCITTTCAA 1026	

RESULT 4
ABS54149
ID ABS54149 standard; cDNA; 828 BP.
XX
AC ABS54149;

XX	25-NOV-2002	(first entry)
XX	DT	
XX	DE	Human cDNA encoding a liver tumour marker protein, CRG-Ll.
XX	DE	Human; ss; gene; liver cancer; liver tumour; CRG-Ll;
KW	KW	hepatocellular cancer; chromosome 9p.
XX	OS	Homo sapiens.
OS	XX	
FT	Key	Location/Qualifiers
EH	CDS	1..828
FT	/tag=	a
FT	/product=	"Liver tumour marker protein, CRG-Ll"
XX	US2002115094-A1.	
PN	XX	
PD	22-AUG-2002.	
XX	14-DEC-2001;	2001US-00017410.
PF	14-DEC-2000;	2000US-0255674P.
XX	(FARN/) FARNHAM P J.	
XX	(GRAV/) GRAVEEL C R.	
PA	Farnham PJ,	Graveel CR;
XX	WPI;	2002-706409/76.
XX	P-PSDB;	ABG32881.
DR	Novel polypeptide designated as CRG-Ll,	useful as diagnostic marker for
XX	Liver cancer,	is differentially expressed in liver tumors relative to
PT	normal liver tissues.	
PT	Claim 2;	Page 8; lipp; English.
XX	The invention relates to a polypeptide designated as CRG-Ll), which is	differentially expressed in liver tumours relative to normal expression
CC	in normal liver tissues, designated CRG-Ll. Also included are the	encoding polynucleotides (in the case of the human sequence, mapping to
CC	chromosome 9p), expression constructs, host cells, anti-CRG-Ll	antibodies, identifying modulators of CRG-Ll, and the use of the CRG-Ll
CC	sequence in the diagnosis of hepatocellular cancer in tumour cells from a	liver of a human or non-human animal. The CRG-Ll protein and
CC	polynucleotide are useful as diagnostic markers for a liver cancer in	humans and non-human animals, and as a system for assessing putative
CC	therapeutic agents. The present sequence encodes human CRG-Ll	
XX	Sequence 828 BP;	157 A; 216 C; 211 G; 244 T; 0 U; 0 Other;
SQ	Query Match	15.9%; Score 664.8; DB 6; Length 828;
	Best Local Similarity	87.7%; Pred No. 4.7e-173;
	Matches 726;	Conservative 0; Mismatches 102; Indels 0; Gaps 0
QY	35	ATGGCGCGCCGCACACTGTGTGGAGACCACCTGGCGGGTGGCAGTTCCGAGGTGGATTGGTGTC 94
Db	1	ATGGCGCGCCGCACACTGTGTGGAGACCACCTGCAGGTGGTAGCTCGGAGGTGGACTGGTGC 60
QY	95	GAGGACAACCTACACTTCGTGCCTGCCATTCGCCAGTTCCTACAAACCATCAGCAACGTC 154
Db	61	GAGGACAACCTACACCATCGTGCTCTCTATCGCCGAGTTCCTACACACCATCAGCAANTGTC 120
QY	155	TGTGTTTTTCATTTTACCTCCCACCTCTGCATGTGCTTGTTCGCCAGTACGCAACGTCGTC 214
Db	121	TTAATTTTTCATTTTACCGCCCATCTGCATGTGCTTGTGATGATGATGCAACATGCTTG 180
QY	215	AACAGCGCATCTACTTAATATGGAAGCTCCCTAGTTGTAGTGGGATTTGGATCTGCTCAC 274
Db	181	AACAGTGCATCTACTTAATCTGGACTCTTTTGGTTGTAGTGGGAATTGGATCCGCTCAC 240
QY	275	TTCCCATGCAACGCTCAGATTTTCCCTGGGTCCAGATGCTTTGATGAACCTTGCCATTCCTGGGTT 334

XX SQ Sequence 1623 BP; 305 A; 457 C; 382 G; 479 T; 0 U; 0 Other;

Query Match 10.2%; Score 424.2; DB 10; Length 1623;
Best Local Similarity 81.2%; Pred. No. 3.3e-106;
Matches 506; Conservative 0; Mismatches 113; Indels 4; Gaps 1;

QY 356 TGGTTTCCAGGAGGATTTTACCAAGATCTTTCGGAATGACAGGGCGAGGTTCAAGGCA 415
DB 920 TTGTTTCGTCAAGGATCTTACCAAGATCTTTCGGAATGACAGGGCGAGGTTCAAGGTT 979

QY 416 GTGGTGTGTCCTGCTGCAATTAACAAGCTGCTGCGCTTTTATCAAGCCCGCATCAAC 475
DB 980 GTGGTGTGTCCTGCTGCAATTAACAAGCTGCTGCGCTTTTATCAAGCCCGCATCAAC 1039

QY 476 AATATTTCCCTGATGATTTTCGGGACCTTCCATGACCTGCGTGTGTTGTCAGAGCTGAAG 535
DB 1040 AACATCTCTGATGACCTTGGGAGTTCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099

QY 536 AGGTGTGACAAATGTGGTGTGTTTAAAGCTGGGCTCTTCTCTGGCTCTGCTGCTGCTGCTG 595
DB 1100 AGGTGTGACAAATGTGGTGTGTTTAAAGCTGGGCTCTTCTCTGGGCTCTGCTGCTGCTG 1159

QY 596 GCTCTCTCTGCTGATCAGCAGCAAGCCTTCTGAGCTGCTCTCTCTCTCTCTCTCTCTCT 655
DB 1160 GCCCTGTTCTGCTGATGACCTGAGTACCGAGCTTCTGCGAGCTGCTGCTGCTGCTGCT 1219

QY 656 CCTTACCTGCACTGTGTGGCATATTTCTCATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 715
DB 1220 CCTTACCTGCACTGTGTGGCATATTTCTCATCTGCTTCTGCTGCTGCTGCTGCTGCTG 1279

QY 716 TGCTTCCCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
DB 1280 TGCTTCCCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1339

QY 776 TGGCCAGCGAGAAATGGCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
DB 1340 TGGCCAGCGAGAAATGGCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1399

QY 836 AAGAAGTCGCCAGTCAAGATCAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTGCA 895
DB 1400 AAGAAGTCAATCAGTCAAGATCAGTGTGCAAGGAGTGTGCAAGGAGTGTGCAAGGAGTGTG 1459

QY 896 CTATTCGAGTGGCGCTGGCTTCTGTTGCTAGCAAGATGCTGAGGGGCTTGAGGAAT 955
DB 1460 CCCCTC---ATGCAGTGGGCTTCTTCTGCTAGGAAGACAGCCAGGGAGTTCGAATAGT 1515

QY 956 TGGTGTGCTGGTGTGTTTAA 978
DB 1516 TGGGTGAGGGCTATGTTTAA 1538

RESULT 7
ID ABK62212/c
XX ABK62212 standard; cDNA; 595 BP.
AC ABK62212;
XX
DT 18-JUN-2002 (first entry)
XX
DE Rat sequence differentially expressed in response to a hepatotoxin #119.
XX
XX Rat; se; hepatotoxin; expressed sequence tag; EST; drug screening;
KW differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
XX
PN WO200210453-A2.
XX
XX 07-FEB-2002.
PD
XX
PF 30-JUL-2001; 2001WO-US023872.
XX

PR 31-JUL-2000; 2000US-0222040P.
PR 02-NOV-2000; 2000US-0244880P.
PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX WPI; 2002-241625/29.
XX
XX Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or cells
PT exposed to the toxin and comparing these to gene expression in unexposed
PT tissues or cells.
XX
PS Claim 1; SEQ ID NO 119; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic effect
CC of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression. The
CC method can also be used to identify an agent which modulates the toxic
CC response and predict cellular pathways that a compound modulates in a
CC cell. The methods utilize a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridises to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity is
CC characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent
XX
SQ Sequence 595 BP; 197 A; 101 C; 146 G; 151 T; 0 U; 0 Other;

Query Match 9.6%; Score 400.4; DB 6; Length 595;
Best Local Similarity 87.9%; Pred. No. 7.2e-100;
Matches 531; Conservative 0; Mismatches 56; Indels 17; Gaps 8;

QY 3557 ATTCTATTGGTGGATTTCTCAGGAGTCTGAAATCTTCCCTCAGAGTCTCTCTCCCAA 3616
DB 595 ATTCTATTGGTGGATTTCTCAGGAGTCTGAAATCTTCCCTCAGAGTCTCTCTCCCAA 536

QY 3617 CCCCTACCCAGAGCAACACTGGGATTTGAACCTTC---TCTGACTCTCTTCTTCCCTC 3672
DB 535 CCGCTCCAGATC---CACTGGGATTTGAACCTTCAAATCTGCTCTCTTCTTCTTCTTC 479

QY 3673 AGGTCTGACAACTAATGGTCTCTGGGGACACCAAGCTAGGGCTTCCCAATCTCTT-AT 3731
DB 478 AGGTCTGACAACTAATGGTCTCTGGGGACACCAAGCTAGGGCTTCCCAATCTCTTCTTCT 419

QY 3732 CCAGCTGAACTGGATTTCTTCCCAAGGGCTTGACCTGGGTGCTGTGGTCCCACTGG 3791

Db 418 CCAGCCAACTGGATCTTCCACACAGGCTTGACCTGGGCTGTTGGTCCCACTGC 359
QY 3792 CCACAAACATCTTTGGCCAGATGGGATCTCAATAGATTTTATAGACATTAATCTCC 3851
Db 358 CCAATGACACATCTTTGACAGATGGGATCTCAATAGATTTTATAGACATTAATCTCC 299
QY 3852 CACAGACTTT--AAACATGGCTGTGTCTTCCATACACATCCGGTCAGATTTAAACT 3909
Db 298 CACAGACTTTTAAACAAACATGGGCTGTGTCTTTCTTCTACACA---AGTCAGATTTAAACT 242
QY 3910 ATTTTATAACACAGCAATTAACCAAGCAATAGAGTACTTTTCAGATATATAACTGTGT 3969
Db 241 ATTTTATACTACAATAATCAACCAAGCAATAGAGTACTTTTCAGATATATGC--CAAT 184
QY 3970 TCATACTTTATGTAGAGTGTGTATGTATAGCGGTATGTACCTGGCTGAAGTAATAT 4029
Db 183 TTATACTTTATGTAGAGTGTGTATGTACAGGTGGCATGTACCTGGCTGAGTAACAT 124
QY 4030 AACCATAGCTCTGGGAGGATTTACAGACTTTTGGCACTTTATGC--TTTTTGTGAACCTCT 4088
Db 123 AGTCATGTCTGGG--GGATTTACAGATCTTTTGGCACTTTATGCTTTTTTGTGAACCTCT 65
QY 4089 GATAACCATGTGTCAATTAATAAGCCCAATAAATGGCAATTTTCTGTGAATAAACAATTTATGT 4148
Db 64 GATAACCATGTGTCAATTAATAAGCCCAATAAATGGCAATTTTCTGTGAATAAACAATTTATGT 5
QY 4149 GTAT 4152
Db 4 ATCT 1

RESULT 8

ADB55226/c
ID ADB55226 standard; DNA; 595 BP.
AC ADB55226;
DT 04-DEC-2003 (first entry)
DE Toxicity-related gene, SEQ ID 252.
KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW drug screening; toxicity assay; ds.
XX Unidentified.
XX WO2003064624-A2.
XX 07-AUG-2003.
XX 31-JAN-2003; 2003WO-US003194.
XX 31-JAN-2002; 2002US-00060087.
XX 15-MAR-2002; 2002US-0364045P.
XX 15-MAR-2002; 2002US-0364055P.
XX 30-DEC-2002; 2002US-0436643P.
XX (GENE-) GENE LOGIC INC.
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elshoff M;
XX WPI; 2003-689530/65.
XX Predicting a toxic effect of a compound, useful in identifying toxicity
XX markers in liver tissues or cells for drug screening and toxicity assays,
XX comprises preparing gene expression profile of tissue or cells exposed to
XX the compound.
XX Claim 1; SEQ ID NO 252; 1156pp; English.
XX The present invention relates to a method for predicting a toxic effect
XX of a compound. The method comprises preparing a gene expression profile

CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 595 BP; 197 A; 101 C; 146 G; 151 T; 0 U; 0 Other;

Query Match 9.6%; Score 400.4; DB 10; Length 595;
Best Local Similarity 87.9%; Pred. No. 7.2e-100;
Matches 531; Conservative 0; Mismatches 56; Indels 17; Gaps 8;

QY 3557 ATTCTATTGGTTGGATTTCTCAGAGTCTGAATCTTCCTCAGAGTCTCTTCTCCCAA 3616
Db 595 ATTCTATTGGTTGGATTTCTCAGAGTCTGAATCTTCCTCAGAGTCTCTTCTCCCAA 536
QY 3617 CCCCTACCCAGAGCCAAACACTGGGATTTGAACCTTC---TCTGACTCTTCTTCCCTC 3672
Db 535 CGCCTCCGATC---CACTGGGATTTGAACCTTCGAATCTTCTTCTTCTTCTTCTC 479
QY 3673 AGGTCTGACAACTAATGTCTCTGGGGACACCCAGCTAGGGCTTCCCAACTCTT-AT 3731
Db 478 AGGTCTGACCACTAAGGCTCTCTGGGGACACCTAAGCTAGGGCTTCCCAACTCTTCT 419
QY 3732 CCAGCTGAACTTGGATTTCTTCCCAACAGGGCTTGACCTGGGTGTGTGGTCCCACTGG 3791
Db 418 CCAGCCAACTTTGGATTTCTTCCCAACAGGGCTTGACCTGGGTGTGTGGTCCCACTGC 359
QY 3792 CCAACAAACATCTTTGGCCAGATTTGGGATTTCTCAATAGATTTTATAGACATTAATCTCC 3851
Db 358 CCAATGACACATCTTTGGCCAGATTTGGGATTTCTCAATAGATTTTATAGACATTAATCTCC 299
QY 3852 CACAGACTTT--AAACATGGCTGTGTCTTTCCATACACATCCGGTCAGATTTAAACT 3909
Db 298 CACAGACTTTTAAACAAACATGGGCTGTGTCTTTCTTCTTCTTCTTCTTCTTCTTCT 242
QY 3910 ATTTTATAACACAGCAATTAACCAAGCAATAGAGTACTTTTCAGATATATAACTGTGT 3969
Db 241 ATTTTATACTACAATAATTAACCAAGCAATAGAGTACTTTTCAGATATATATGC--CAAT 184
QY 3970 TCATACTTTATGTAGAGTGTGTATGTATAGCGGTATGTACCTGGCTGAGTAATAT 4029
Db 183 TTATACTTTATGTAGAGTGTGTATGTACAGGTGGCATGTACCTGGCTGAGTAATAT 124
QY 4030 AACATAGCTCTGGGAGATTTACAGACTTTTGCATTTATGC--TTTTTGTGAACCTCT 4088
Db 123 AGTCATGTCTGGG--GGATTTACAGATCTTTTGGCACTTTTATGCTTTTTTGTGAACCTCT 65
QY 4089 GATAACCATGTCAATTAATAAGCCCAATAAATGGCAATTTTCTGTGAATAAACAATCATAT 4148
Db 64 GATAACCATGTCAATTAATAAGCCCAATAAATGGCAATTTTCTGTGAATAAACAATTTATGT 5
QY 4149 GTAT 4152
Db 4 ATCT 1

RESULT 9

ADB55226/c
ID ADB55226 standard; cDNA; 1527 BP.
XX
XX AC ADB55226;
XX

DT

06-MAY-2004

(first entry)

XX

Human 3T3 cell conversion promoter PPI1646 cDNA.

XX

3T3 cell conversion; promoter; human; gene; ss.

XX

Homo sapiens.

XX

CNI403477-A.

XX

19-MAR-2003.

XX

12-SEP-2001; 2001CN-00126725.

XX

12-SEP-2001; 2001CN-00126725.

XX

(SHAN-) SHANGHAI XINSHIJI GENE TECHN DEV CO LTD.

XX

Gu J, Yang S;

XX

WPI; 2003-494226/47.

DR

P-PSDB; ADL06641.

XX

Human protein with function of promoting 3T3 cell conversion and its coding sequence.

XX

Example 1; SEQ ID NO 1; 41pp; Chinese.

XX

This invention describes a novel human protein with 3T3 cell conversion promoting function, polynucleotides encoding the polypeptide and the recombinant process of producing the polypeptide. The present invention also discloses the agonist resisting the polypeptide and its treatment effect. The present invention also discloses the application of the polynucleotides encoding the human protein with 3T3 cell conversion promoting function.

XX

Sequence 1527 BP; 376 A; 357 C; 383 G; 411 T; 0 U; 0 Other;

XX

Query Match

XX

Best Local Similarity

XX

Matches 475; Conservative

XX

0; Mismatches 74; Indels 97; Gaps 1;

QY

142

GATCAGCAAGCTCTTTTTCATTTTACCTCCCATCTGCATGCTTGTTCGCGCAGTA

201

Db

178

GATCAGCAAGCTCTTTTTCATTTTACCTCCCATCTGCATGCTTGTTCGCGCAGTA

237

QY

202

CGCAACGTCGTTTCAACAGCGGCATCTACTTAATATGACGCTCTCTAGTTAGTGGGAT

261

Db

238

TGCAACATGCTTCAACAGTGGCATCTACTTAATCTGGACTCTTTTGGTTGTAGTGGGAT

297

QY

262

TGGATCTGTCTACTTCCATGCAACGCTGATTTCTGGTTCAGATGCTTGATGAACCTTGC

321

Db

298

TGGATCCGCTACTTCCATGCAACCTTAGTTTCTTGGTTCAGATGCTTGATGAACCTTGC

357

QY

322

CATTCTGCGGTTCTGATGTGCTTTGGCCATGTGTTTCCAGAGAGTATTACCAAA

381

Db

358

AGTCCTTTGGGTTCTGATGTGCTTTGGCCATGTGTTTCCCAAGAGTATACCAAA

417

QY

382

GATCTTTCCGAATGACAGCGGCAGTTCAAGCAGTGTGTGTCCTGTCTGCAATATAC

441

Db

418

GATCTTTCCGAATGACCGGGTAGTTCAAGGTGGTGGTCACTGCTCTGCGGTAC

477

QY

442

AACGTGCTTGGCTTTATCAAGCCGCCATCAACATATTTCCCTGATGATTTCTGGACT

501

Db

478

GACGTGCGCTGGCATTTGTCAAGCCTGCCATCAACACATCTCTCTGATGACCTGGAGT

537

QY

502

TCCATGCACTGCGCTGCTTTGTCAGAGCTGAAG-----

535

Db

538

TGCTTGCACTGCACTGCTCATGCAGAGCTTAAGAGGATGAGAGAACCGAGGAG

597

QY

536

535

Db

598

ACACAGGAAGGTGGCCAGCAAGAGGTGGAGACAAGGTCTGACGATGATGACTCTCTG

657

QY

536

AGGTGTGACAAATGCTGCTGTGTTTAAAGCTGGGCTCTTCTCTGGCCCTCT

584

Db

658

GACCCCGTGCAGGTGTGACAAATGCTGTGTTTAAAGCTGGGCTCTTCTCTGGCCCTCT

717

QY

585

GGTGGACTCTGGCTCTCTTCTGCTGATCAGCGCAACAGCCTTCTGTGAGCTGTCTCTCT

644

Db

718

GGTGGACCTGGCCCTCTTCTGCTGATCAGTGACCGAGCTTCTGCGAGCTGTCTCAT

777

QY

645

CCTTTTCACTTCCCTTACCTGCTGCTGTGTGGCATATTTCTCATCTG

690

Db

778

CCTTCAACTTCCCTTACCTGCTGCTGCTGTGTGGCATATTTCTCATCTG

823

RESULT 10

ADL06642

ADL06642

standard; DNA; 1527 BP.

XX

ADL06642;

XX

06-MAY-2004

(first entry)

DT

Human 3T3 cell conversion promoter PPI1646 DNA.

DE

3T3 cell conversion; promoter; human; gene; ds.

XX

Homo sapiens.

XX

Key

Location/Qualifiers

FT

CDS

218..640

FT

/*tag= a

FT

/product= "PPI1646"

XX

CNI403477-A.

XX

19-MAR-2003.

XX

12-SEP-2001; 2001CN-00126725.

XX

12-SEP-2001; 2001CN-00126725.

XX

(SHAN-) SHANGHAI XINSHIJI GENE TECHN DEV CO LTD.

PI

Gu J, Yang S;

XX

WPI; 2003-494226/47.

DR

P-PSDB; ADL06641.

XX

Human protein with function of promoting 3T3 cell conversion and its coding sequence.

PT

Example 1; SEQ ID NO 3; 41pp; Chinese.

XX

This invention describes a novel human protein with 3T3 cell conversion promoting function, polynucleotides encoding the polypeptide and the recombinant process of producing the polypeptide. The present invention also discloses the agonist resisting the polypeptide and its treatment effect. The present invention also discloses the application of the polynucleotides encoding the human protein with 3T3 cell conversion promoting function.

CC

Sequence 1527 BP; 376 A; 357 C; 383 G; 411 T; 0 U; 0 Other;

XX

Query Match

XX

Best Local Similarity

XX

Matches 475; Conservative

XX

0; Mismatches 74; Indels 97; Gaps 1;

QY

142

GATCAGCAAGCTCTTTTTCATTTTACCTCCCATCTGCATGCTTGTTCGCGCAGTA

201

Db

178

GATCAGCAAGCTCTTTTTCATTTTACCTCCCATCTGCATGCTTGTTCGCGCAGTA

237

QY

202

CGCAACGTCGTTTCAACAGCGGCATCTACTTAATATGACGCTCTCTAGTTAGTGGGAT

261

Db

238

TGCAACATGCTTCAACAGTGGCATCTACTTAATCTGGACTCTTTTGGTTGTAGTGGGAT

297

QY

262

TGGATCTGTCTACTTCCATGCAACGCTGATTTCTGGTTCAGATGCTTGATGAACCTTGC

321

Db

298

TGGATCCGCTACTTCCATGCAACCTTAGTTTCTTGGTTCAGATGCTTGATGAACCTTGC

357

QY

322

CATTCTGCGGTTCTGATGTGCTTTGGCCATGTGTTTCCAGAGAGTATTACCAAA

381

Db

358

AGTCCTTTGGGTTCTGATGTGCTTTGGCCATGTGTTTCCCAAGAGTATACCAAA

417

QY

382

GATCTTTCCGAATGACAGCGGCAGTTCAAGCAGTGTGTGTCCTGTCTGCAATATAC

441

Db

418

GATCTTTCCGAATGACCGGGTAGTTCAAGGTGGTGGTCACTGCTCTGCGGTAC

477

QY

442

AACGTGCTTGGCTTTATCAAGCCGCCATCAACATATTTCCCTGATGATTTCTGGACT

501

Db

478

GACGTGCGCTGGCATTTGTCAAGCCTGCCATCAACACATCTCTCTGATGACCTGGAGT

537

QY

502

TCCATGCACTGCGCTGCTTTGTCAGAGCTGAAG-----

535

Db

538

TGCTTGCACTGCACTGCTCATGCAGAGCTTAAGAGGATGAGAGAACCGAGGAG

597

QY

536

535

Db

598

ACACAGGAAGGTGGCCAGCAAGAGGTGGAGACAAGGTCTGACGATGATGACTCTCTG

657

DR 238 TGCAACATGCTTCAACAGTGGCATCTACTTAATCTGGACTCTTTGGTTGTAGTGGAAAT 297
XX
DR 262 TGGATCTGCTACTTCCATGCAAGCTGAGTTTCTGGGTGATGCTTGATGAATTCG 321
XX
PT TGGATCCGCTACTTCCATGCAAGCTGAGTTTCTGGGTGATGCTTGATGAATTCG 357
PT
PT 322 CATTCCTGTGGGTCTGATGCTGCTTGGCCATGCTTCCAGAGGTATTTACCAA 381
XX
XX 358 AGTCCTTTGGGTCTGATGCTGCTTGGCCATGCTTCCAGAGGTATTTACCAA 417
PS
XX 382 GATCTTTCGGAATGACAGGGCAGGTTCAAGCAGTGTGTGCTCTGTCGAATAC 441
CC
CC 418 GATCTTTCGGAATGACAGGGGTAGGTTCAAGGTGTGTGCTCACTGCTCTGCGTTAC 477
CC
CC 442 AACGTGCTTGGCTTATCAAGCCGCCATCAACAATATTTCCCTGATGATTTCTGGACT 501
CC
CC 478 GACGTGCTTGGCTTATCAAGCCGCCATCAACAATATTTCTGATGACCTGGGACT 537
CC
CC 502 TCCATGCACTGGCTGCTTGTTCAGAGCTGAAG----- 535
CC
CC 538 TCCTTGCACTGCACTGCTCATCGCAGAGTAAAGAGGCATGAGAGAACCGAAGGAG 597
CC
CC 536 ----- 535
DB 598 ACACAGAAAGGTGGCCAGCAGGAGGTGGAGACAAGGTCTGACGATGAGTCTCTG 657
DB
QY 536 -----AGGTGTCACATGTCGCTGTTTAAAGTGGGCTCTTCTGCGCTCT 584
QY
DB 558 GACCCCGTGGAGGTGACAACTGCTGTTTAAAGTGGGCTCTTCTGCGGCTCT 717
DB
QY 585 GGTGACTCTGGCTCTTCTGCTGATCAGCGACCAAGCCTTCTGAGCTGCTCTCT 644
QY
DB 718 GGTGACCTGGCTCTTCTGCTGATCAGTACCGAGCTTTCGCGAGCTGCTGTCT 777
QY
DB 645 CCTTCACTCCCTTACTGCACTGTGTGTGGCATATTTCTCATCTG 690
DB
DB 778 CTTTCAACTTCCCTTACTGCACTGATGCTGCTGAGTGGTGGACGTCAGATG 823

RESULT 11

ID AD121468 standard; cDNA; 487 BP.
XX
AC AD121468;
XX
DT 15-APR-2004 (first entry)
XX
DE Novel human expressed sequence tag, EST #167.
XX
KW forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;
KW ss; Gene; EST; expressed sequence tag.
OS Homo sapiens.
XX
XX WO2003025148-A2.
XX
XX 27-MAR-2003.
XX
XX 19-SEP-2002; 2002WO-US029964.
XX
XX 19-SEP-2001; 2001US-0323739P.
XX
XX 13-SEP-2002; 2002US-00323739.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue Au, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX

DR
DR
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ

Query Match
Best Local Similarity 7.1%; Score 294.8; DB 10; Length 487;
Matches 317; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 10 GCTGCTCCGATGCTCCAGAGCGGCATGGGCGCCCGCAGCTGTGGGACCACTGGCGGC 69
DB 134 GCTGCTCCGATGCTCCAGAGCGGCATGGGCGCCCGCAGCTGTGGGACCACTGGCGGC 193
QY 70 TGGCAGTTCCGAGTGGATTTGGTGGAGGACAACTACACTATCGTCCCTGCAATTCGCGA 129
DB 194 TGGTAGCTCGGAGTGGACTGGTGGAGGACAACTACACTATCGTCCCTGCAATTCGCGA 253
QY 130 GTTCTCAACACAGATCAGCAGCTGTTGTTTTCATTTTACCTCCCATCTGCGATGCTTCT 189
DB 254 GTTCTCAACACAGATCAGCAGCTGTTGTTTTCATTTTACCTCCCATCTGCGATGCTTCT 313
QY 190 GTTCCGCGAGTACGCAACAGCTGCTTCAACAGCGGATCTACTTAATATGGAGCTCCTAGT 249
DB 314 GTTTCGTGATGCAACATGCTTCAACAGTGGGATCTACTTAATCTGGATCTTTTGGT 373
QY 250 TGTAGTGGGATTTGGATCTGCTACTTCCATGCAACGCTGAGTTTCTTGGGTTCAGATGCT 309
DB 374 TGTAGTGGGAAATTTGGATCCGCTACTTCCATGCAACCTTAGTTTCTTGGGTTCAGATGCT 433
QY 310 TGATGAACTTGGCAATTTCTGTTGGTCTCTGATGCTGCTTGGCCATGCTGTTTCC 363
DB 434 TGATGAACTTGGCAATTTCTGTTGGTCTCTGATGCTGCTTGGCCATGCTGTTTCC 487

RESULT 12

AAK51566
ID AAK51566 standard; cDNA; 1215 BP.
XX
AC AAK51566;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 111.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.

WPI; 2003-354603/33.
P-P8DB; ADI21688.
New polynucleotides and secreted proteins, useful for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve tissue growth or regeneration, in wound healing, and in tissue repair and replacement.
Example 2; SEQ ID NO 719; 156bp; English.
The invention relates to an isolated polynucleotide encoding a polypeptide with biological activity. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders and other traits, to assess biodiversity, as nutritional sources or supplements. The polynucleotides may also be used as molecular weight markers, chromosome markers or map related gene positions, or as an antigen to raise anti-DNA antibodies or elicit immune responses. The polypeptides are useful for raising antibodies, as markers for tissues in which the corresponding polypeptide is expressed, for re-engineering damaged or diseased tissues, for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and/or nerve tissue growth or regeneration, in wound healing, in tissue repair and replacement, in healing of burns, incisions and ulcers, and in treating cancer. The present sequence represents a novel human expressed sequence tag, EST.
Sequence 487 BP; 81 A; 138 C; 131 G; 134 T; 0 U; 3 Other;

XX WO200157190-A2.
FN XX
PD XX
PI XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX PF
XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
XX P-PSDB; AAM78433.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 1; Page 750-751; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
XX SQ Sequence 1215 BP; 213 A; 374 C; 309 G; 319 T; 0 U; 0 Other;
Query Match 4.2%; Score 176.8; DB 4; Length 1215;
Best Local Similarity 51.3%; Pred. No. 9.9e-38;
Matches 437; Conservative 0; Mismatches 412; Indels 3; Gaps 1;
Qy 74 AGTTCCGAGGTGGATTGGTGGGAGGACAACTACACTATCGTCTGCCATTCGCGAGTTC 133
Db 80 AGCTCCGAGGTGGACTGGTGTGAGGCAACTTCCAGTACTCGGAGCTGGTGGCGAGTTC 139
Qy 134 TACAACACGATCAGCAAGCTCTGTTTTCATTTTACCTCCATCTGCATGTGCTTGTTC 193
Db 140 TACAACAGTTCTCAATATCCCTTCTTATCTTCGGGCCACATGATGATGCTCTGATG 199
Qy 194 CGCCAGTACGCAAGCTGCTTCAACAGCGGCACTACTTAATATGAGAGCTCCTAGTTGTA 253
Db 200 CACCGGTATGCCAGAGCGCTCCCGCTACATTTTACGTTGTCTGGTCTCTTCATGATC 259
Qy 254 GTGGGATTTGATCTGTCTACTTCCATGCAAGCTGATTTTCCTGGGTCCAGATCCTGAT 313
Db 260 ATAGGCTGTGTTCTCCATGATATTTCCATATGACGCTCAGCTTCTGGGCCAGCTGCTGAC 319
Qy 314 GAATCTGCACTCTGTGGTGTCTGATGTGCTTTGGCCATGTGTTTCCAGGAGGTAT 373
Db 320 GAGATCGCCATCCTGTGGCTCTTGGGCACTGGCTATAGCATATGATGATCCCGCTGCTAT 379
Qy 374 TTACAAAGATCTTTCCGAATGACAGGGGCGAGTTCAAGGCGAGTGTGTGTCTGTCT 433
Db 380 TTCCCTCTCTTCTTGGGGGGAACAGGTCCCGAGTTTCACTCCGCTGTGCTTCTATCACCAC 439

QY 434 GCAATTACAAAGTGTGCTTGGCGTTTATCAAGCCGCGCATCAACAAATATTTCCCTGATGATT 493
Db 440 GTGGTCAAGCACT 499
QY 494 CTGGGACTTTCCATGCACTGCGCTGCTGTTGTGAGAGCTGAAGAGGTGTGACAAATGTGCGT 553
Db 500 ATTGGCCCTGCACATTCTCTACATCGTGTGCCAGGAGTACAGGAAGACCAAGCAATAAGGAG 559
QY 554 GTGTTTAAGCTGGGCT 613
Db 560 CTTTGGCACCTGATGAGGTCT 619
QY 614 AGCGACCAAGCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 673
Db 620 AGTGACCGTCTGCTTTCAGCTTCTGCGAGAGGATTCATTTCTTCTATCTGCACAGATC 679
QY 674 TGGCATATTTCTCATCTGCGCTTGTCTGCTACCTGGGCTGTGTGTGCTTTCGCTCTCTTTGAT 733
Db 680 TGGCATGTGCTCATCAGCATCACCTTCCCTTATGCGATGCTCACCATGGCTTGGTGGAT 739
QY 734 GCTGCTCAGAGATACCTGAGCAAGGTCCAGTATCATGATTTCTGGCCAGCGAGAAATGG 793
Db 740 GCCAACTATGAGATGCCAGGTGAACCCCTCAAAGTCCGCTACTTGGCTTCGGGACAGTTGG 799
QY 794 GCTTTTATTTGCT 853
Db 800 CC---CGTGGGCTGCGCTACGTGGAAATCCGGGGTGTGACAAAGGACTGCTGAGACCTG 856
QY 854 ATCAGGTGATGGCAAGGAGTACAGCTTCTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTG 913
Db 857 CCAGCTCTTGTGATATCAACACCAACCACTTGCCTGTGTCTTGTGAGAAAGATAGCCCCG 916
QY 914 GCCTTCGTTTTC 925
Db 917 GTCAGGACTTTC 928
RESULT 13
AAK52550
ID AAK52550 standard; cDNA; 1631 BP.
XX
XX AAK52550;
XX
XX 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 2079.
DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation; ss.
XX Homo sapiens.
XX OS
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX PF
XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR P-PSDB; AAM79417.
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX Claim 1; Page 4462; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 1631 BP; 247 A; 399 C; 344 G; 350 T; 0 U; 291 Other;

Query Match 4.2%; Score 176.8; DB 4; Length 1631;
Best Local Similarity 51.3%; Pred. No. 1.2e-37;
Matches 437; Conservative 0; Mismatches 412; Indels 3; Gaps 1;

QY 74 AGTTCGAGGTGATGCTGGAGGACAACTACACTATGTCGCTGCCATGCGAGTTC 133
DB 80 AGTTCGAGGTGATGCTGGAGGACAACTACACTATGTCGCTGCCATGCGAGTTC 139

QY 134 TACAACACGATCAGACGCTCTGTTTTCATTTTACCTCCATCTGCATGCTGCTTC 193
DB 140 TACAACAGTTCCTCAATATCCCTCTTCACTTCGGGCGACATGATGCTCCTGATG 199

QY 194 CGCCAGTACGCAAGCTGCTTCAACAGCGGATCTACTTAATATGAGCGCTCTAGTTGTA 253
DB 200 CACCGGTATGCCAGAGCGCTCCGCTACATTTACGTTGCTGGGCTCTTTCATGATC 259

QY 254 GTGGGATTCGATGCTCTACTTCCATGCAACGCTGATTTCTGGGTGATGCTGAT 313
DB 260 ATAGGCTGTTCTCATGATTTCCACATGACGCTCAGCTTCTGGGCGAGTCTGGAC 319

QY 314 GAACCTTCCCATCTGTCGGTCTGATGCTGCTTGGCCATGTTGCTCCAGGAGTAT 373
DB 320 GAGATCGCCATCTGTGGCTCTGGGCGATGGCTATAGCATATGATGCCCGCTGCTAT 379

QY 374 TTACCAAGATCTTTTGGAAATGACAGGGGAGGTTCAAGGCGATGTTGTCCTGCT 433
DB 380 TTCCCTCTCTCTTCTGGGGGAAACAGGTCACAGTTTCACTCGCCTGGTCTTCATCACCAC 439

QY 434 GCAATTACACGCTGCTGGGTTTATCAAGCCGCCATCAACATATTTCCCTGATGAT 493
DB 440 GTGGTACGACCTTCTGCTCTCTCTGGGCGCCACGCTCAACGCTACGCTCTCAACAGC 499

QY 494 CTGGGACTTCCATGCACTGCGCTGCTTGTGTCAGAGCTGAAGAGGTGTGACATGTCGT 553
DB 500 ATTGCCCTGCACATCTCTACATCGTGTGCCAGGATGACAGGACCAATTAAGAG 559

QY 554 GTGTTTAAGCTGGGCTCTTCTGCGCTCTGGTGGATCTGGCTCTCTCTGCTGGATC 613
DB 560 CTTCGGCACCTGATGAGGCTCGTGGTGTATGAGGCTGTGCTCTGACACGCTGATC 619

QY 614 AGCGACCAAGCTTCTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 673
DB 620 AGTGACCGTCTGCTTGTGACGCTTCTGCGAGAGGATTCATTTCTTCTATCTGCGACGATC 679

QY 674 TGGCATATCTCATCTGCTCTGCTTCTGATCCTGGGCTGCTGCTGCTGCTGCTGCTGCT 733

DB 680 TGGCATGCTCTATCAGCATCACCTTCCCTTATGGCATGGTCAACATGGCCCTTGGTGGAT 739
QY 734 GCTGCTTCAGAGATACCTGAGCAAGTCCAGTTCATCAGATTCTGGCCAGCGAGAAATGG 793
DB 740 GCCAACTATGAGATGCCAGGTGAACCTCAAAGTCCGCTACTGCTCTGGGACAGTTGG 799
QY 794 GCTTTTATGCTGCTCCCTTATGCTGCTCTGTCGCCCAAGAAAGTCCGCTGCTGAG 853
DB 800 CC---CGTGGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
QY 854 ATCAGGTGATGGCAAGGAGTACGACCTTCTCTACTTACTTCTTCTGAGTGGCGCTG 913
DB 857 CCAGCTCTTGACTATCCACCACCCCAACAACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
QY 914 GGCTTCGTTTGC 925
DB 917 GTCAGGACTTGC 928

RESULT 14
AAH48639
ID AAH48639 standard; DNA; 792 BP.
XX
AC AAH48639;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human ceramidase K1 DNA.
XX
KW Ceramidase; human; K1; antiproliferative; anticancer; anti-eczema;
KW antiporiasis; dermatological; ceramide; sphingosine; treatment; ds;
KW altered cell proliferation; cancer; neurodermatitis; eczema; cosmetic;
KW permeability barrier; psoriasis; ichthyosis; lamellar ichthyosis ICR2.
XX
OS Homo sapiens.
XX
PN WO200155408-A1.
XX
PD 02-AUG-2001.
XX
PF 27-JAN-2001; 2001WO-EP000900.
XX
PR 27-JAN-2000; 2000DE-01003293.
PR 09-MAR-2000; 2000DE-01011392.
XX
PA (MEMO-) MEMOREC MEDICAL MOLECULAR RES COLOGNE ST.
XX
PI Hofmann K, Conradt M;
XX
DR WPI; 2001-483256/52.
DR P-PSDB; AAB86363.
XX
XX New ceramidase containing specific structural motif, useful for
PT diagnosis, prevention and treatment of ceramidase defects, e.g.
PT ichthyosis, also in cosmetics.
XX
XX Claim 3; Page 16; 31pp; German.
CC
CC This invention describes novel human ceramidase (I) containing a specific
CC structural motif which has antiproliferative, anticancer, anti-eczema,
CC antiporiasis and dermatological activity. (I) cleave ceramide, resulting
CC in formation of sphingosine. (I), or the nucleic acid (II) encoding it,
CC are used for diagnosis, prevention or treatment of diseases associated
CC with ceramide defects, particularly altered cell proliferation (cancer)
CC or altered ceramide layers on the skin (neurodermatitis, eczema,
CC psoriasis), also for targeted modification of the permeability barrier by
CC ceramidase or its activators, e.g. for transcutaneous delivery of
CC substances. Specifically they are used for diagnosis of ichthyosis,
CC particularly lamellar ichthyosis ICR2. (I) can also be used in cosmetics.
CC This sequence encodes human ceramidase K1
XX
SQ Sequence 792 BP; 145 A; 241 C; 200 G; 206 T; 0 U; 0 Other;

Query Match	4.2%	Score 174.8;	DB 4;	Length 792;
Best Local Similarity	52.6%;	Pred. No. 2.8e-37;		
Matches	380;	Conservative 0;	Mismatches 342;	Indels 0; Gaps 444
74	AGTTTCGAGGTGGATTGGTCGCGAGGACAACTACACTATCGTGCCTGCCTACCTATTCGCCAGTTCCCGAGTTC	133		
25	AGCTCCGAGGTGAGCTGGTGTGAGAGCACTTCAGTACTCGGAGCTGGTGCCGAGTTC	84		
134	TACAACAAGATCAGCAACAGCTCTTTGTTTTCATTTTACCTCCCATCTGCATGTGCTGTGTTTC	193		
85	TACAACAAGCTTCTCCAATATCCCTCTTTCATCTTCGGGCCACTGATGATGCTCCTGATG	144		
194	GGCCAGTACCGAAGTGCCTTCAACAGGGGCACTACTTAATATGAGAGCTCCTAGTTGTA	253		
145	CACCCGTATCCCAAGAGCGCTCCCGCTACATTTACGTTGTCTGGGTCTCTTTCATGATC	204		
254	GTGGGGATTGGATCTGTCTACTTCCATGCAAGCTGAGTTTCTCGGGTCAGATGCTTGAT	313		
205	ATAGCGCTGTTCTCCATGATTTTCCATGACGCTCAGCTTCTGCGGCAGCTGCTGGAC	264		
314	GAACTTCCCATCTGTGGGTTCATGATGTGTGCTTTGGCCATGTGGTTTCCGAGAGGTAT	373		
265	GAGATGCCATCTGTGGCTCTCTGGGCAGTGGCTATAGCATATGAGATGGATGCCCGCTGCTAT	324		
374	TTACCAAGATCTTTCGGATGACAGGGGAGGTTCAGGCGAGTGTGTGTCCTGTCT	433		
325	TTCCCTCTCTCTTGGGGGGAACAGTCCCAAGTTTCATCGGCTGTGCTTCAACCACT	384		
434	GCAATTAAACAGTCTTGGCGTTTATCAAGCCGCCATCAACAATATTTCCCTGATGATT	493		
385	GTGCTCAGCACCCCTTCTGTCCTTCTGCGGCCCAAGCTCAAGCCCTCAACAGC	444		
494	CTGGGACTTCATGCACTGGCTGCTTGTGTGAGAGCTGAGAGGTGTGACATGTGGCT	553		
445	ATTCCCTTGACATTCCTCTACATCGTGTGCAGGAGTACAGGAAGACCAGCAATAAGGAG	504		
554	GTGTTTAAAGCTGGCCCTCTTCTGCGCTCTCGTGTGACTCTGGCTCTCTTCTGTGTGATC	613		
505	CTTCGGCACCTGATGAGGTCTCGTGGTTTATGGGTGTGCTCTGACCAAGCTGGATC	564		
614	AGCGACCAAGCCTTCTGTGAGTGTCTCTCTCTCTTCACTTCCCTTACCTGCACTGTGTG	673		
565	AGTGACCGTCTGCTTTTGCAGTCTTGGCAGAGGATTCATTTCTTCTATCTGCACAGCATC	624		
674	TGGCATATTCATCTGCCTTGTCTGTGTAAGTGGGTGTGTGCTTGGCTACTTTGAT	733		
625	TGGCATGTGTCTATCAGCATCACTTCCCTTATGGCATGTGTCCCATGGCTTGGTGGAT	684		
734	GCTGCCCTCAGAGATPACTCTGAGCAAGGTCCAGTCAATCAGATTCTGGGCCACGAGAAATGG	793		
685	GCCAACTATGAGATGCCAGGTGAAACCCCTCAAAGTCCGCTACTTGGCTCTGGGACAGTTGG	744		
794	GC	795		
745	CC	746		

RESULT 15	
ABL24043	
ID ABL24043 standard; DNA; 537 BP.	
XX	
AC ABL24043;	
XX AC	
XX DT	
XX 26-MAR-2002 (first entry)	
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23602.	
XX	
XX	
KW Drosophila; developmental biology; cell signalling; insecticide;	
XX pharmaceutical; gene; ds.	
XX	
XX Drosophila melanogaster.	
OS	
XX	
PN WO200171042-A2.	
QY	353 ATGTGGTTCCAGGAGGTATTATCCAAAGATCTTTCGGAATGACAGGGGAGGTTCAAG 412
Db	349 CTCCTCTATCGAAGCGATACTATCCCAAGTTCGTGAAACGATCGAAACCTTCAGT 408
QY	413 GCAGTGTGTGTCTGTCTGCAATTACAAAGTGTCTTGGGTTTATCAAGCCCGGCATC 472
Db	409 TGGCTCATGCTCTTGTCTGGCGAATTGCTCGACGCGCTTGTCTGGTGTGAAGCCCAATTGTT 468
QY	473 AACAAATATTTCCCTGATGATCTCGGACCTTCCATGCACCTGGCTGCTTGTTCAGAGCTG 533
Db	469 AACCCCTTTGTTCTCATGAGTGTGCCACCATGTTAATGCTCTACACAGAGCTG 529
QY	533 AAGAGGT 539
Db	529 CAGAGGT 535

Search completed: November 20, 2004, 21:28:44
Job time : 1777.81 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:51:27 ; Search time 357.166 Seconds
(without alignments)
8308.591 Million cell updates/sec

Title: US-10-017-410-1

Perfect score: 4175
Sequence: 1 ggcacagggtgctcgat.....aaaaaaaaaaaaaaaaaac 4175

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	114.2	2.7	358	4	US-09-621-976-11241
2	55.8	1.3	7218	1	US-08-232-463-14
3	46	1.1	1141	4	US-09-806-708B-22
4	45.4	1.1	1063	4	US-09-799-451-875
5	44.8	1.1	832	4	US-09-621-976-2813
6	42.4	1.0	832	4	US-09-621-976-2813
7	41.8	1.0	30310	4	US-09-657-346A-96
8	41	1.0	289	3	US-09-007-005-17
9	41	1.0	289	3	US-09-244-796-17
10	40.8	1.0	435	4	US-09-248-796A-6940
11	40	1.0	1044	4	US-09-495-406-12
12	40	1.0	1044	4	US-09-816-028A-16
13	40	1.0	1044	4	US-10-303-162-16
14	40	1.0	11474	4	US-09-495-406-1
15	40	1.0	11474	4	US-09-816-028A-1
16	40	1.0	11474	4	US-10-303-162-1
17	39.2	0.9	1141	4	US-09-806-708B-22
18	38.8	0.9	302	3	US-09-328-111-679
19	38.4	0.9	2044	4	US-09-327-138C-38
20	38.4	0.9	2863	4	US-09-327-138C-2
21	38.4	0.9	3646	4	US-09-327-138C-36
22	38.4	0.9	3665	4	US-09-327-138C-35
23	38.4	0.9	3686	4	US-09-327-138C-3
24	38.4	0.9	3686	4	US-09-327-138C-37
25	38.4	0.9	3940	4	US-09-327-138C-1
26	38.2	0.9	168174	4	US-10-071-411A-63
27	38.2	0.9	168273	4	US-10-071-411A-2

c 28	37.8	0.9	1821	4	US-09-149-476-90	Sequence 90, Appl
29	37.8	0.9	48974	3	US-08-920-422-17	Sequence 17, Appl
30	37.6	0.9	604	4	US-09-401-064-126	Sequence 126, App
31	37.6	0.9	640681	4	US-09-790-988-1	Sequence 1, Appli
32	37.4	0.9	1131	2	US-08-951-924A-1	Sequence 3, Appli
33	37.4	0.9	1131	3	US-09-172-339-3	Sequence 3, Appli
34	37.4	0.9	1131	3	US-09-420-211-10	Sequence 10, Appl
c 35	37.4	0.9	90050	3	US-09-245-041-5	Sequence 5, Appli
c 36	37.4	0.9	90050	4	US-09-358-055B-5	Sequence 5, Appli
c 37	37.4	0.9	90050	4	US-09-893-238-5	Sequence 9223, Ap
c 38	37.2	0.9	239	4	US-09-513-999C-9223	Sequence 45, Appl
39	37	0.9	596	1	US-08-150-331-45	Sequence 45, Appl
40	37	0.9	596	4	US-08-569-284-45	Sequence 45, Appl
c 41	37	0.9	691	4	US-09-270-767-29557	Sequence 29557, A
c 42	37	0.9	1356	4	US-09-270-767-13559	Sequence 13559, A
43	37	0.9	5134	2	US-08-310-912A-157	Sequence 157, App
44	37	0.9	5134	3	US-09-301-085-157	Sequence 157, App
45	37	0.9	5134	5	PCT-US95-04589-157	Sequence 157, App

ALIGNMENTS

RESULT 1
US-09-621-976-11241/c
; Sequence 11241, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; FILE OF INVENTION: ESTS and Encoded Human Proteins.
; TITLE OF INVENTION: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.Pm
; SEQ ID NO 11241
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-11241

Query Match	2.7%;	Score 114.2;	DB 4;	Length 358;
Best Local Similarity	90.4%;	Pred. No. 2.1e-25;		
Matches 122;	Conservative	0;	Mismatches 13;	Indels 0; Gaps 0;
Qy	10	GCTCTCGATGCTCCAGAGCGCCATGGCGCCCGCCGCTGTGGGACACCTGCGGGC	69	
Db	136	GCTGCTCAATGGCCCGAGTGGCCATGGGCGCCCGCCTGGTGGGACACCTGCGGGC	77	
Qy	70	TGGCAGTTTCGAGGTGGATTGGTCGAGGAGCACTACACTATCGTGCCTGCTTCCGA	129	
Db	76	TGGTAGCTCGAGGTGGACTGGTGGAGGAGCACTACACCATCGTGCCTGCTTCCGA	17	
Qy	130	GTTCTCAACACGAT 144		
Db	16	GTTCTCAACACGAT 2		

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F1s
US-08-232-463-14

Query Match 1.3%; Score 55.8; DB 1; Length 7218;
Best Local Similarity 8.1%; Pred. No. 8.3e-06;
Matches 36; Conservative 222; Mismatches 189; Indels 0; Gaps 0;
QY 480 TTTCCTGATGATTCGGGACTTCCAGTCCGCTGCTGTGTGAGAGCTGAAGAGGT 539
DB 988 TATAACTATTTTCCTCGTTGGCCATAGCTCAGAGATTAATTCGAGCTGGCTGCA 1047
QY 540 GTGACATGCTGCTGTTAAGCTGGCTCTCTCTGGCTCTGCTGCTGCTGCTGCTC 599
DB 1048 GTGAGGAGCTTGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1107
QY 600 TCTTCTGCTGATCAGCAGCAGCCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCT 659
DB 1108 YY 1167
QY 660 ACTGCACTGTGTGGCATATTCATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 719
DB 1168 YY 1227
QY 720 TCGCTACTTGTGCTGCTCAGATACCTGAGCAAGTCCAGTCAATCAGATTCGCTG 779
DB 1228 YY 1287
QY 780 CCAGCAGAAATGGGCTTTATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 839
DB 1288 YY 1347
QY 840 AGTCGCCAGTCAAGTACGATGCGCAGGAGGAGTCAAGCTTCTCTACTTACTTCTAT 899
DB 1348 YY 1407
QY 900 TCGAGTGGCTGGCTGCTGCTGCT 926
DB 1408 YY 1434

RESULT 3

US-09-806-708B-22
Sequence 22, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1141)
OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 1.1%; Score 46; DB 4; Length 1141;
Best Local Similarity 9.9%; Pred. No. 0.0022;
Matches 75; Conservative 280; Mismatches 405; Indels 1; Gaps 1;
QY 2135 TCTAGGATCTTACTCTCTGTTTAGAACCTTTCACAGGGGTACAAAGTGGGAAGTGAACCTTAA 2194
DB 115 TNKMGKTGWRHRYWRWMBEDTVDHVVVTAMNNAWTTCMMMDKDKRTRWWKNNNATG 174
QY 2195 AGAGTTTAAATGAATAGACTCTCTGCTTTTGGCTTTGGTGGAGTTTAAAGTAAC 2254
DB 175 WDDDTKYHMNNNNCBVTWVRYKTRDRWSBRNNYGBWKNWSDVYVYVWVDDMC 234
QY 2255 TTGCTGGCTATTCTCTGTTTATAGTCACAAATATTTATAGACATGAAGTGTAAATA 2314
DB 235 KKKRVWRVTRGRMYVMBTAHRRRYNNGWTBAMAYRRWTMNNNNNNNAKAMCKRAKY 294
QY 2315 AGTTGTCTTTTATTAATTCATAGCATTTACCAACTCTCCAGGTAGCAACACACAGCTC 2374
DB 295 WGNRABVNSTCTTKVTSKWANNCRAGDANKDKHKKWKSAAAGVYNNNNNN 354
QY 2375 ATTTGATTCGGATTTTACAGGCAAGAACTACTTGGTTTGAATGATGTATATGCTATTT 2434
DB 355 WTYKARHBAARDWVWHSAAKWHANAHAHYSRKKWTBYRKTWVNNNNNGTTMMKRWAWY 414
QY 2435 TACTGTAGCTTAAATATATGTTTAAA-ATGACTTTGAGTGAATGTTTGTGGCAACCTAGG 2493
DB 415 WKMDMDWBTYNNNNNGRTYYGWTNKKRWYTKWANNCKWRANDHKTCNNNTTWK 474
QY 2494 GGTATTGATCAGAAATTCGGCTGAGCTCTTTGGTTTGTAGTCTAGAGTTTTCAGAG 2553
DB 475 MKTYNNCYKSWNTGKSHRBAAAVVTWMMWRRYAHANNNNNDWYWKACTWYKYBVCS 534
QY 2554 GCGAAATCAAACTAGCTTGTCTGTTGTTTATCTTACAGGCTCGGAGCTCCCTTAT 2613
DB 535 KWNMYAAWYTKSSWNTYSRYRWKTNNNSWRNRSDTRSMGRANNYARABHYKYNTRW 594
QY 2614 ATATTCTACATAGGACTTTATTATACATAAGTCTGTATAATGTCCTCGAAGATGACACT 2673
DB 595 BMSHTWEHBRAGAHHWMBMYBAKCHMKAWYKAKYAGAGSNNNNNNNNNNNNNN 654
QY 2674 AGCTGCTTTTCTGGAAGGCTGCTGGGCTGGAGAGTTGGTTCAACAGTTAAGAAC 2733
DB 655 ATCARDYAAAGRWYAMANAkWYKBAANNAYYTHANNWGWNNATDTRRTMKNNN 714
QY 2734 CGTGTGCTCTCTCCAGAGCAGCCAGCTGCTGAGCTCTATACAGGACTCCAGC 2793
DB 715 NNAGTWKNNNNNNNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAK 774
QY 2794 CCTGGGCACTCTGAGCCCTCTTCTGACTTCTTAAGGGCATTTGGTGCACACGCTACACA 2853


```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
; US-09-621-976-2813

Query Match 1.0%; Score 42.4; DB 4; Length 832;
Best Local Similarity 12.7%; Pred. No. 0.023;
Matches 43; Conservative 156; Mismatches 137; Indels 2; Gaps 1;

QY 2085 TTCTGAGTGTGCTGCTGAGCTGAGGAGCTAGCTGGGGTCTTGGCACTTCTAGGATCT 2144
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
360 WTWKTKWYWTYTTWTWMMWKKKRRWYWKSTYACSRKYRTWGWMMWKKRWMT 301
QY 2145 CTTACTCTGTTAGAACCTTCACAGGTACAAAGTGGGAAGTGAAGTAAAGATTTT 2204
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 RWCYCMKWCYRGRCAWYTWARGWWSYANGKWSMSAMCTRMYYKKGSTYWTM 241
QY 2205 AATGAATAGACTTCATCTGCTTTGTGGCTTTCGAGTCTTAAAGTAACCTGCTGGCT 2264
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
240 KCTCAWYCYWYKWKRWKSKTCSGSRGGYMTSYSTRSYWYASWMTYMCWNGRW 181
QY 2265 ATTCTTGTGTTATAGTCACAAATATTTATAGAACATGAAGTGTAAATAAGTTGCTTT 2324
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 STYWMANGKWKRYATTWRRAMWMAWMTMMWMMWCMSSRGAAYRRTM--MMWG 123
QY 2325 TATTAATTCATAGCATTTACCACTCTCCAGGTAGCAACACACAGTCAATTTGATGG 2384
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 YRYWKKSYRRTRCAWAYAKTKRSYYWCMRWKWKRCMMWMMWMAWYKTMWRAWCKT 63
QY 2385 CGATTTACAAGGCGAGGAATACTGCTTTTGAATGATTG 2422
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 RYRWNAWAWRWMTMMWYTWRAKRRWWRKWR 25

RESULT 7
US-09-657-346A-96
; Sequence 96, Application US/09657346A
; Patent No. 6503754
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; TITLE OF INVENTION: ANTISENSE MODULATION OF BH3 INTERACTING DOMAIN DEATH AGONIST
; TITLE OF INVENTION: EXPRESSION
; FILE REFERENCE: RFS-0135
; CURRENT APPLICATION NUMBER: US/09/657,346A
; CURRENT FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 174
; SEQ ID NO 96
; LENGTH: 30310
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19791)...(19802)
; NAME/KEY: CDS
; LOCATION: (21160)...(21370)
; NAME/KEY: CDS
; LOCATION: (24168)...(24307)
; NAME/KEY: CDS
; LOCATION: (25696)...(25908)
; NAME/KEY: CDS
; LOCATION: (27235)...(27246)
; US-09-657-346A-96

Query Match 1.0%; Score 41.8; DB 4; Length 30310;
Best Local Similarity 71.4%; Pred. No. 0.76;
Matches 55; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 2781 CAGCAACTCCAGCCCTCGGGGCACTCTGAGCCCTCTTCTGACTTCTAAGGGCAATTTGTGC 2840
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
11806 CTGTAATCCAGTTTCCAGTTCATCTGGCACCCCTCTTCTGGCTCCACGGGCACCGAGAC 11865
QY 2841 ACAGTACATACAGGCA 2857
```

```
Db 11866 AGAGATACATGCAGGCA 11882

RESULT 8
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
; US-09-007-005-17

Query Match 1.0%; Score 41; DB 3; Length 289;
Best Local Similarity 7.6%; Pred. No. 0.026;
Matches 17; Conservative 97; Mismatches 111; Indels 0; Gaps 0;

QY 541 TGACAAATGTCGTGTGTTAAAGCTGGGCCTCTTCTCTGGCCTCTGTGGACTCTGGCTCT 600
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
233 YGYAYGYCYTYGYSYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 174
QY 601 CTTCGTCTGATCAGGACCAAGCCCTCTCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
173 YNYSYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 114
QY 661 CTGCACTGTGTGTGGCATATTCTCATCTGCTTCTGTCTGTCTGTCTGTCTGTCTGTCT 720
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
113 YNYSYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 54
QY 721 CGCCTACTTTGATGCTGCTCCCTCAGAGATACCTGAGCAAGGTCAGT 765
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
53 YTYGYTAYAVTYTYGYTAYAVTYGYTAYAVTYGYTAYAVTYGYTAYAVTYGY 9

RESULT 9
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
```

```
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match      1.0%; Score 41; DB 3; Length 289;
Best Local Similarity 7.6%; Pred. No. 0.026;
Matches 17; Conservative 97; Mismatches 111; Indels 0; Gaps 0;

QY 541 TGCAATGTGCGTGTGTTAAGCTGGGCTCTTCTCTGGCCCTCTGGTGACTCTGGCTCT 600
DB 233 YGYCAYGCTGTGYGYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 174
QY 601 CTTCTGCTGGATCAGCGACCAAGCCTTCTGTCAGCTGCTCTCTCTCTCTCTCCCTA 660
DB 173 YNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 114
QY 661 CTTGACCTGTGTGGCATAATCTCATCTGCTCTGCTCTGCTGCTGGCTGTGCTGT 720
DB 113 YNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 54
QY 721 CGCCTACTTTGATGCTGCTCAGATACCTGAGCAAGTCCAGT 765
DB 53 YTYTYGYTAYATYTYGYTAYAYATYAYATYAYATYTYGY 9

RESULT 10
US-09-248-796A-6940/c
; Sequence 6940, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6940
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (387),(391)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-248-796A-6940

Query Match      1.0%; Score 40.8; DB 4; Length 435;
Best Local Similarity 47.3%; Pred. No. 0.042;
Matches 123; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 2251 TAACCTGTCTGGCTATTCTTGTATTATAGTCACAAATTTATTAGAACATGAAGGTGAA 2310
DB 330 TGAATAGCTGTGTGTGTGTATTATTATTATTATTATTATTATTATTATTATTATTATT 271
QY 2311 AATAAGTTGTCTTTTATTAAATTCATAGCATTTACCACTCTCCAGGTAGCAACAC 2370

; EARLIER APPLICATION NUMBER: 09/495,406
; Sequence 12, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-0001100S
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1044)
; OTHER INFORMATION: beta-1,4 N-acetylglactosaminyl (GalNAc)
; OTHER INFORMATION: transferase from C. jejuni strain OH4384 (ORF 5a
; OTHER INFORMATION: of LOS biosynthesis locus)
US-09-495-406-12

Query Match      1.0%; Score 40; DB 4; Length 1044;
Best Local Similarity 45.7%; Pred. No. 0.16;
Matches 139; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 3862 AAAACATGGCTTGTGTCTTTCCATACACATCGGTCAGATTAAACTATTTTATAACCA 3921
DB 963 ATAGAAATTTTAAATTTCTTCAATCTTATCAATTTATTTATTAATTTCACTGTAATATCTGCT 904
QY 3922 CAGGAATTTAAACCAAGCAATAGAGTACTTTTCAGATATAAACTGTGTTTCATCTTTATG 3981
DB 903 ATGTTTTTTTAAAGGAAATGANAATTTATTTAATTTCACTGTAATATATCTGCTG 844
QY 3982 TAGAGTGTCTATGTATAGGGCGGTATGTACCTCGCTGAAGTAATTAATTAACCATAGCTCT 4041
DB 843 TTTATACTTTAAATTCCTCTAAAGAGTGTATATTTGAAGAAAACCTTGTGTTGTTCTT 784
QY 4042 GGGAGGATTACAGACCTTTTGGCACTTTATGCTTTTTTGTGAACTGTGATAACCATGTC 4101
DB 783 TTTCCAATTTGTTATTGTTCTGTCAATTAATPACTTCTTCCATTTTAAGTTATAATTATT 724
QY 4102 AATATTTAAAGCAATACTGGCATTTTCTGTGAATAAACATGCATATGTATCTAATAAAA 4161
DB 723 GATAAGCAAGCAATCTCTGTTCTTTCAATATATGCTCATTTATTTTACAAAGAAA 664
QY 4162 AAAA 4165
DB 663 AAAA 660
```

RESULT 12

US-09-816-028A-16/c
; Sequence 16, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1044)
; OTHER INFORMATION: beta-1,4 N-acetylgalactosaminyl (GalNAc)
; OTHER INFORMATION: transferase from C. jejuni strain OH4384 (ORF 5a)
; OTHER INFORMATION: of lipooligosaccharide (LOS) biosynthesis locus
US-09-816-028A-16

Query Match 1.0%; Score 40; DB 4; Length 1044;

Best Local Similarity 45.7%; Pred. No. 0.16; Indels 0; Gaps 0;
Matches 139; Conservative 0; Mismatches 165;
Qy 3862 AAAACATGGCTTGTCTCTTCCATACACATCCGGTCAGATTAAACCTATTTTATAACCA 3921
Db 963 ATAGAAATTTTAAATTTCTTCAATCTATCCCAATATATTTATAAATATCTTGAGCTCT 904
Qy 3922 CAGGAATTAACCAAGCAATAGAGTACTTTCAGATATAAACTGTGTTTCATCTTTATG 3981
Db 903 ATGTTTTTTTTAAAAAGGAAATGATAATTTATTTAATTCAGTGTGAATAATATCCTGTG 844
Qy 3982 TAGAGTGTCTATGTATAGGGGTATGACCTCGCTGAAGTATATTAACCATAGCTCT 4041
Db 843 TTTATACCTTTAATTTGCTCTAAGAGTGTATTTTGAAGAAATACTTTGTTTGTGCTTT 784
Qy 4042 GGGAGGATTTACAGACCTTTTGCACCTTTATGCTTTTGTGAACCTCTGATAACCATGGTC 4101
Db 783 TTTCCAAATTTGTTATTTGATTTCTGCAATTAATTAATTTTAAAGTTATAATTATT 724
Qy 4102 AATATTAAGCAATAACTGCAATTTCTGTGAATAAATGCAATGATATGATCTAAAAAAA 4161
Db 723 GATAAGCAAGCAATCTCCTGTTCTTCTCAATATATATGCAATTTATTTTACAAAGAAA 664
Qy 4162 AAAA 4165
Db 663 AAAA 660

RESULT 13

US-10-303-162-16/c
; Sequence 16, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1044)
; OTHER INFORMATION: beta-1,4 N-acetylgalactosaminyl (GalNAc)
; OTHER INFORMATION: transferase from C. jejuni strain OH4384 (ORF 5a)
; OTHER INFORMATION: of lipooligosaccharide (LOS) biosynthesis locus
US-10-303-162-16

Query Match 1.0%; Score 40; DB 4; Length 1044;

Best Local Similarity 45.7%; Pred. No. 0.16; Indels 0; Gaps 0;
Matches 139; Conservative 0; Mismatches 165;
Qy 3862 AAAACATGGCTTGTCTCTTCCATACACATCCGGTCAGATTAAACCTATTTTATAACCA 3921
Db 963 ATAGAAATTTTAAATTTCTTCAATCTATCCCAATATATTTATAAATATCTTGAGCTCT 904
Qy 3922 CAGGAATTAACCAAGCAATAGAGTACTTTCAGATATAAACTGTGTTTCATCTTTATG 3981
Db 903 ATGTTTTTTTTAAAAAGGAAATGATAATTTATTTAATTCAGTGTGAATAATATCCTGTG 844
Qy 3982 TAGAGTGTCTATGTATAGGGGTATGACCTCGCTGAAGTATATTAACCATAGCTCT 4041
Db 843 TTTATACCTTTAATTTGCTCTAAGAGTGTATTTTGAAGAAATACTTTGTTTGTGCTTT 784
Qy 4042 GGGAGGATTTACAGACCTTTTGCACCTTTATGCTTTTGTGAACCTCTGATAACCATGGTC 4101
Db 783 TTTCCAAATTTGTTATTTGATTTCTGCAATTAATTAATTTTAAAGTTATAATTATT 724
Qy 4102 AATATTAAGCAATAACTGCAATTTCTGTGAATAAATGCAATGATATGATCTAAAAAAA 4161
Db 723 GATAAGCAAGCAATCTCCTGTTCTTCTCAATATATATGCAATTTATTTTACAAAGAAA 664
Qy 4162 AAAA 4165
Db 663 AAAA 660

RESULT 14

US-09-495-406-1/c
; Sequence 1, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:53:12 ; Search time 1950.22 Seconds
(without alignments)
11567.962 Million cell updates/sec

Title: US-10-017-410-1

Perfect score: 4175
Sequence: 1 ggcacgagggtgtccgat.....aaaaaaaaaaaaaac 4175

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4175	100.0	4175	13	US-10-017-410-1
2	723.8	17.3	4202	16	US-10-188-832-21
3	723.8	17.3	4212	15	US-10-295-027-151
4	664.8	15.9	828	13	US-10-017-410-3
5	653.8	15.7	822	15	US-10-182-447-6
6	400.4	9.6	595	9	US-09-917-800A-119
7	329.2	7.9	744	10	US-09-945-527-48
8	174.8	4.2	792	15	US-10-182-447-4
9	131.2	3.1	1098	15	US-10-172-118-2618
10	131.2	3.1	1098	16	US-10-342-887-2618
11	125.8	3.0	1473	15	US-10-106-698-611
12	65	1.6	418	16	US-10-242-535A-50197

13	1.6	418	16	US-10-085-783A-50197	Sequence 50197, A
14	1.5	77530	17	US-10-322-281-265	Sequence 265, App
15	1.5	659158	9	US-09-771-208-20	Sequence 20, Appl
16	1.4	35425	15	US-10-017-161-2429	Sequence 2429, Ap
17	1.4	35425	15	US-10-292-798-2069	Sequence 2069, Ap
18	1.3	91609	17	US-10-322-696-109	Sequence 109, App
19	1.3	19619	15	US-10-369-214-129	Sequence 129, App
20	1.2	7516	15	US-10-311-455-34	Sequence 34, Appl
21	1.2	102374	13	US-10-087-192-667	Sequence 667, App
22	1.2	50295	11	US-09-997-722-359	Sequence 259, App
23	1.2	115780	17	US-10-367-094-96	Sequence 96, Appl
24	1.2	60968	15	US-10-085-117-181	Sequence 181, App
25	1.1	52479	11	US-09-997-722-61	Sequence 61, Appl
26	1.1	97415	17	US-10-322-281-287	Sequence 287, App
27	1.1	50657	17	US-10-322-281-601	Sequence 601, App
28	1.1	215980	10	US-09-972-546-16	Sequence 16, Appl
29	1.1	32495	17	US-10-322-281-401	Sequence 401, App
30	1.1	67191	11	US-09-997-722-169	Sequence 169, App
31	1.1	67191	15	US-10-105-612-1	Sequence 1, Appl
32	1.1	185555	13	US-10-087-192-1999	Sequence 1999, Ap
33	1.1	171936	15	US-10-265-071-24	Sequence 24, Appl
34	1.1	171936	15	US-10-025-966A-24	Sequence 24, Appl
35	1.1	73507	13	US-10-087-192-1147	Sequence 1147, Ap
36	1.1	96593	11	US-09-997-722-151	Sequence 151, App
37	1.1	197775	13	US-10-087-192-853	Sequence 853, App
38	1.1	38342	16	US-10-221-714A-472	Sequence 472, App
39	1.1	54648	15	US-10-085-117-289	Sequence 289, App
40	1.1	801	15	US-10-182-447-5	Sequence 5, Appl
41	1.1	1063	16	US-10-302-172-875	Sequence 875, App
42	1.1	10279	15	US-10-311-455-1563	Sequence 1563, Ap
43	1.1	10279	16	US-10-240-589C-85	Sequence 85, Appl
44	1.1	55906	17	US-10-322-281-417	Sequence 417, App
45	1.1	305	9	US-09-967-769A-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-10-017-410-1
; Sequence 1, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Graveel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; CURRENT FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4175
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(859)
US-10-017-410-1

Query Match	100.0%;	Score 4175;	DB 13;	Length 4175;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4175;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GGCAGAGGGCTCCCGATGCTCCAGAGCGGCCCATGGCGCCCGCACTGGTGGACCA	60	
Db	1	GGCAGAGGGCTCCCGATGCTCCAGAGCGGCCCATGGCGCCCGCACTGGTGGACCA	60	
Qy	61	CCTGCGGGCTGCAGTTCGAGGTGGATGGTGGAGGACAACTACATCTATCGTGGCTGC	120	
Db	61	CCTGCGGGCTGCAGTTCGAGGTGGATGGTGGAGGACAACTACATCTATCGTGGCTGC	120	
Qy	121	CATTGCGAGTCTTACACACGATCAGCAACGCTCTTTGTTTTTACCTTCCCATCTG	180	

121	Db	 CATTGCCAGTCTTACACACGATCAGCAACGCTCTTCTTTTTCATTTTTACCTCCCATCTG	180
181	Qy	CATGTGCTTGTTCGCGCAGTACGCAACGTCGTTTCAACAGCGGCATCTACTTAATATGGAC	240
181	Db	CATGTGCTTGTTCGCGCAGTACGCAACGTCGTTTCAACAGCGGCATCTACTTAATATGGAC	240
241	Qy	GCTCCTAGTTGTAGTGGGGATTGGATCTGTCTACTTCATCGCAACGCTGAGTTTCCCTGGG	300
241	Db	GCTCCTAGTTGTAGTGGGGATTGGATCTGTCTACTTCATCGCAACGCTGAGTTTCCCTGGG	300
301	Qy	TCAGATGCTTGATGAACHTTGCATTCTGTGTGGTTTCTGATGTGTGCTTTTGGCCATGTGGTT	360
301	Db	TCAGATGCTTGATGAACHTTGCATTCTGTGTGGTTTCTGATGTGTGCTTTTGGCCATGTGGTT	360
361	Qy	TCCCAGAGAGTATTTTACCAAGATCTTTTCGGAATGACAGCGGCGCAGGTTCGAAGCAGTGT	420
361	Db	TCCCAGAGAGTATTTTACCAAGATCTTTTCGGAATGACAGCGGCGCAGGTTCGAAGCAGTGT	420
421	Qy	GTGTGTCTGTCTCAATTACAAAGTGTCTTGGCGTTTATCAAGCCCGCCATCAACAATAT	480
421	Db	GTGTGTCTGTCTCAATTACAAAGTGTCTTGGCGTTTATCAAGCCCGCCATCAACAATAT	480
481	Qy	TTCCCTGATGATTCCTGGGAATTCCATGACATGCGCTGCTTGTTCGAGAGCTGAAGAGGTG	540
481	Db	TTCCCTGATGATTCCTGGGAATTCCATGACATGCGCTGCTTGTTCGAGAGCTGAAGAGGTG	540
541	Qy	TGACAAATGTGCGTGTGTTAAGCTGGGCGCTTCTCTGGCCTCTGGTGTGATCTGGGCTCT	600
541	Db	TGACAAATGTGCGTGTGTTAAGCTGGGCGCTTCTCTGGCCTCTGGTGTGATCTGGGCTCT	600
601	Qy	CTTCTGTGTGATCAGCGACCAAGGCTTTCTGTGAGCTGCTCTCCCTCTTTTCACTTCCCCCTA	660
601	Db	CTTCTGTGTGATCAGCGACCAAGGCTTTCTGTGAGCTGCTCTCCCTCTTTTCACTTCCCCCTA	660
661	Qy	CCTGCACTGTGTGTGGCATATTTCTCATCTGTGCTTGTGTGCTGTGTGCTGTGTGCTCT	720
661	Db	CCTGCACTGTGTGTGGCATATTTCTCATCTGTGCTTGTGTGCTGTGTGCTGTGTGCTCT	720
721	Qy	CGCCTACTTTGATGCTGCGCTCAGAGATACCTGAGCAAGGTCAGTCAATCAGATTCGCGC	780
721	Db	CGCCTACTTTGATGCTGCGCTCAGAGATACCTGAGCAAGGTCAGTCAATCAGATTCGCGC	780
781	Qy	CAGCGAATAATGGGCTTTTATTGGTGTCCCTTATGTGTCCCTTCTGTGTGCCCAACAAGAA	840
781	Db	CAGCGAATAATGGGCTTTTATTGGTGTCCCTTATGTGTCCCTTCTGTGTGCCCAACAAGAA	840
841	Qy	GTGCGCAGTCAAGATCAAGTATCGCAAGGCAAGGAGTGACAGCTTCTCTACTTTCTTCTATT	900
841	Db	GTGCGCAGTCAAGATCAAGTATCGCAAGGCAAGGAGTGACAGCTTCTCTACTTTCTTCTATT	900
901	Qy	CGAGTGGCGCTGGGCTTCGTTTGTCTAGCAAGATGGCTGAGGGGTTGAGGAATGGTG	960
901	Db	CGAGTGGCGCTGGGCTTCGTTTGTCTAGCAAGATGGCTGAGGGGTTGAGGAATGGTG	960
961	Qy	TGTTGTGGGTGTTTTAAAATTCTGCTTCTTTGTGTATCTAACTGGACCACTGTGCTGACCT	1020
961	Db	TGTTGTGGGTGTTTTAAAATTCTGCTTCTTTGTGTATCTAACTGGACCACTGTGCTGACCT	1020
1021	Qy	CCCTAGGTTTAAGTATAGGGCTCAGGGACATAAGGTGTCTTCCCTCAGTATCCTTTCCAGGC	1080
1021	Db	CCCTAGGTTTAAGTATAGGGCTCAGGGACATAAGGTGTCTTCCCTCAGTATCCTTTCCAGGC	1080
1081	Qy	ATACGGGCTGTCTGGGTTATGTCCTAATGACATCAACAGATGATGTTCTTTGGGAGACCT	1140
1081	Db	ATACGGGCTGTCTGGGTTATGTCCTAATGACATCAACAGATGATGTTCTTTGGGAGACCT	1140
1141	Qy	AGGCGAACCCCAAGTTTCTTGTCTGGAGAGGTTAGCTTGTGTTTCCACCAATTTCCAGACT	1200
1141	Db	AGGCGAACCCCAAGTTTCTTGTCTGGAGAGGTTAGCTTGTGTTTCCACCAATTTCCAGACT	1200
1201	Qy	CTCAACCCCTGTATAAACAGACCTTCTGATTTGGTGATGAAGAGGTTCCAGAACTTTTCA	1260

Db	1201	CTCAACCCCCTGATAAACAAGACCTTCTGTGATTTGGTGTAGAAAGGTTCCAGAACTTTTTC	1261
QY	1261	TTTTGCCGGAGAAACTGTCTCTTCAACAAACCAAGTGGGCAAAACAACGTGTGGGGGGTG	1320
Db	1261	TTTTGCCGGAGAAACTGTCTCTTCAACAAACCAAGTGGGCAAAACAACGTGTGGGGGGTG	1320
QY	1321	TGACTAAGACGGGTGGCTTGTCACTCGCACCTCTTTAGCTTTTCCAGGTTCTCCACGTGTT	1380
Db	1321	TGACTAAGACGGGTGGCTTGTCACTCGCACCTCTTTAGCTTTTCCAGGTTCTCCACGTGTT	1380
QY	1381	TGTGGATCCTTCCACTACTCTGCTGAGAGATGGAGCCACGGCTTCAGAGGGCAAGCTG	1440
Db	1381	TGTGGATCCTTCCACTACTCTGCTGAGAGATGGAGCCACGGCTTCAGAGGGCAAGCTG	1440
QY	1441	GCAACACCTCTATGCCAAAGCTACACTCCTCTTTAAAGCACACATTACATAGACACTAT	1500
Db	1441	GCAACACCTCTATGCCAAAGCTACACTCCTCTTTAAAGCACACATTACATAGACACTAT	1500
QY	1501	TTCTGTCTTTCCAGAGTGCAGCAGCCTCAGACCCACAGAGAATCCTTCAGGTTATGTGAA	1560
Db	1501	TTCTGTCTTTCCAGAGTGCAGCAGCCTCAGACCCACAGAGAATCCTTCAGGTTATGTGAA	1560
QY	1561	GATTCCACACAGACCCCTCTTGTGACCTCTGTGAGAGCAAGCCCTCGTTGTAATCGCG	1620
Db	1561	GATTCCACACAGACCCCTCTTGTGACCTCTGTGAGAGCAAGCCCTCGTTGTAATCGCG	1620
QY	1621	GAGGCACTCTGAGGGCTTCACCTGTGAGCGCTGTGAAGATTTGGAAGCTTTTCTTA	1680
Db	1621	GAGGCACTCTGAGGGCTTCACCTGTGAGCGCTGTGAAGATTTGGAAGCTTTTCTTA	1680
QY	1681	TTGGAAGCTTTGGTGTGTCTGTTACCTCCAAGATTTGACCCCGCTTTCTCTGCTGTTTT	1740
Db	1681	TTGGAAGCTTTGGTGTGTCTGTTACCTCCAAGATTTGACCCCGCTTTCTCTGCTGTTTT	1740
QY	1741	AGGGGATATGTGCTTCTGTAGTGCAGTAGCCTCCCTAGAGATGTGGCCCTCCGGCTTTGT	1800
Db	1741	AGGGGATATGTGCTTCTGTAGTGCAGTAGCCTCCCTAGAGATGTGGCCCTCCGGCTTTGT	1800
QY	1801	TTTTCATAACTCTGGTCTAACTGGTTTCTCAGAGCACTTTGCTCTTTGATGCTGGGC	1860
Db	1801	TTTTCATAACTCTGGTCTAACTGGTTTCTCAGAGCACTTTGCTCTTTGATGCTGGGC	1860
QY	1861	GGTCACTACACTCTGATTTGGCTGGGATCCAGGGAGGAAGGGCGAGACTCTGACA	1920
Db	1861	GGTCACTACACTCTGATTTGGCTGGGATCCAGGGAGGAAGGGCGAGACTCTGACA	1920
QY	1921	TGAATGTCTCACCTGCATCTCACTGTCTTCACTGGGCTCTCTTCAAGGTATGAAGTGG	1980
Db	1921	TGAATGTCTCACCTGCATCTCACTGTCTTCACTGGGCTCTCTTCAAGGTATGAAGTGG	1980
QY	1981	GTATGGGGTATCTCAGGATGTTTGTAACTCAGGCACTTTCTGCTTTCTGACATTCATT	2040
Db	1981	GTATGGGGTATCTCAGGATGTTTGTAACTCAGGCACTTTCTGCTTTCTGACATTCATT	2040
QY	2041	GTCACTGCTGAGACTGCACCTCAGGGACTGGAGTTGGAAACCGCTTCTGAGTGGTCTCT	2100
Db	2041	GTCACTGCTGAGACTGCACCTCAGGGACTGGAGTTGGAAACCGCTTCTGAGTGGTCTCT	2100
QY	2101	GAGTCTGAAGAGCTAGCTCGGGTTCGTGGACTTCTTAGGATCTCTTACTCTGTGTTAGAA	2160
Db	2101	GAGTCTGAAGAGCTAGCTCGGGTTCGTGGACTTCTTAGGATCTCTTACTCTGTGTTAGAA	2160
QY	2161	CCTTCAGAGGTACAAAGTGGGAACCTGACTTAAAGAGTTTTTAAATGAATAGACTTCAT	2220
Db	2161	CCTTCAGAGGTACAAAGTGGGAACCTGACTTAAAGAGTTTTTAAATGAATAGACTTCAT	2220
QY	2221	TCTGCTTTTGTGGCTTTTGGAGTTTTAAAGTAACCTTGCTGGGCTATTTCTGTGTTATAGT	2280
Db	2221	TCTGCTTTTGTGGCTTTTGGAGTTTTAAAGTAACCTTGCTGGGCTATTTCTGTGTTATAGT	2280
QY	2281	CACAAATATTTATAGAACATGAAGGTGTAAAAATAAGTTGTCTTTTATTAATTCATAGCA	2340
Db	2281	CACAAATATTTATAGAACATGAAGGTGTAAAAATAAGTTGTCTTTTATTAATTCATAGCA	2340

QY 2341 TTTACCAACTCTCCAGGTAGCAAAACACACAGTCATTGTGATGTCGATTTTCAAGGCGAG 2400
DB |||||
QY 2341 TTTACCAACTCTCCAGGTAGCAAAACACACAGTCATTGTGATGTCGATTTTCAAGGCGAG 2400
DB |||||
QY 2401 AATACCTGGTTTGAATGATGATATATGTCATTTTACCTGAGCTTAAATATGTTAAAA 2460
DB |||||
QY 2401 AATACCTGGTTTGAATGATGATATATGTCATTTTACCTGAGCTTAAATATGTTAAAA 2460
DB |||||
QY 2461 TGACTTTGAGTGAATGTTTCTGGCAACTAGGGGTTTATGGATCAGAAATGCGGCTGAG 2520
DB |||||
QY 2461 TGACTTTGAGTGAATGTTTCTGGCAACTAGGGGTTTATGGATCAGAAATGCGGCTGAG 2520
DB |||||
QY 2521 TCCTTTGGTTTTGAGTCTAGAAGTTTTCAGAGGCAAAATCAAACTAGCCTGTTGTTCTG 2580
DB |||||
QY 2521 TCCTTTGGTTTTGAGTCTAGAAGTTTTCAGAGGCAAAATCAAACTAGCCTGTTGTTCTG 2580
DB |||||
QY 2581 GTTCATCTTACAGGCTCGGAGCTGCCCTTATATATTTCTACATAGGACTTATATACA 2640
DB |||||
QY 2581 GTTCATCTTACAGGCTCGGAGCTGCCCTTATATATTTCTACATAGGACTTATATACA 2640
DB |||||
QY 2641 TAAGTCTGTATAAATGTCCTGAAGATGACACTAGCTGCCCTTCACTCGAAAGGGTCTGCT 2700
DB |||||
QY 2641 TAAGTCTGTATAAATGTCCTGAAGATGACACTAGCTGCCCTTCACTCGAAAGGGTCTGCT 2700
DB |||||
QY 2701 GGGCTCGAGAGTTGGTTCAACAGTTAAAGAACCCGTTGTTCTCTCCAGACGACCCAG 2760
DB |||||
QY 2701 GGGCTCGAGAGTTGGTTCAACAGTTAAAGAACCCGTTGTTCTCTCCAGACGACCCAG 2760
DB |||||
QY 2761 CGACTTGTAGGCTCATAAACAGCAACTCCAGCCCTGGGCTGATCTGAGCCCTCTTCTGA 2820
DB |||||
QY 2761 CGACTTGTAGGCTCATAAACAGCAACTCCAGCCCTGGGCTGATCTGAGCCCTCTTCTGA 2820
DB |||||
QY 2821 CTTCTAAGGGCATTTGGTGACACCTGACATACAGGCGCAAAACATTTATACATGTAAC 2880
DB |||||
QY 2821 CTTCTAAGGGCATTTGGTGACACCTGACATACAGGCGCAAAACATTTATACATGTAAC 2880
DB |||||
QY 2881 GTAATAAATGATAGTTAGTGAGACGCTGAAGGAAGGAGTTTATAGTCAAGGTTTA 2940
DB |||||
QY 2881 GTAATAAATGATAGTTAGTGAGACGCTGAAGGAAGGAGTTTATAGTCAAGGTTTA 2940
DB |||||
QY 2941 GTCTGACCTGAGTGCCTTCTTACCGGATGGTCTCTCAAGTCAGGGAAGCAAT 3000
DB |||||
QY 2941 GTCTGACCTGAGTGCCTTCTTACCGGATGGTCTCTCAAGTCAGGGAAGCAAT 3000
DB |||||
QY 3001 TTCTTTTTTATCAGGAAAGGACTTATTTCCCTAGGGCCCTCTGTCACATCCCTAGGAACA 3060
DB |||||
QY 3001 TTCTTTTTTATCAGGAAAGGACTTATTTCCCTAGGGCCCTCTGTCACATCCCTAGGAACA 3060
DB |||||
QY 3061 GAGATAAATACGGATGGATGTAATGAACATGCTTGGGTGAAGAGCGGATACCTGA 3120
DB |||||
QY 3061 GAGATAAATACGGATGGATGTAATGAACATGCTTGGGTGAAGAGCGGATACCTGA 3120
DB |||||
QY 3121 CTGGACCCAGTGGGCCCTTCCACAGCGAAGCCCGGTAACCGATGTCACCTCCAGAAATC 3180
DB |||||
QY 3121 CTGGACCCAGTGGGCCCTTCCACAGCGAAGCCCGGTAACCGATGTCACCTCCAGAAATC 3180
DB |||||
QY 3181 TTCTCCCTTCTGGTAGAGATGACATCGATGATGTCGTGACCACTGGGCGCTGCGATG 3240
DB |||||
QY 3181 TTCTCCCTTCTGGTAGAGATGACATCGATGATGTCGTGACCACTGGGCGCTGCGATG 3240
DB |||||
QY 3241 GGTGTGGGCTGTGTTCTGTTCCGAATCTACTGAGATCTAGGATCTAGGACAGAGCAATGA 3300
DB |||||
QY 3241 GGTGTGGGCTGTGTTCTGTTCCGAATCTACTGAGATCTAGGATCTAGGACAGAGCAATGA 3300
DB |||||
QY 3301 ATGTTACCAAGTGTGATGCTGATGATTTTACAGTTTGAACCATTTTGGTTTTT 3360
DB |||||
QY 3301 ATGTTACCAAGTGTGATGCTGATGATTTTACAGTTTGAACCATTTTGGTTTTT 3360
DB |||||
QY 3361 AGGAGAAATCTCTCTCTAGTGCCCTGTGATGCCAAAGCCAGCCCTTCAAGAGTGTTC 3420
DB |||||
QY 3361 AGGAGAAATCTCTCTCTAGTGCCCTGTGATGCCAAAGCCAGCCCTTCAAGAGTGTTC 3420
DB |||||

QY 3421 CTTCTGTCTGCTTCCCTTTGTAATGTGATCCACTCGGGAAATGGTGTCTTACCAGGGA 3480
DB |||||
QY 3421 CTTCTGTCTGCTTCCCTTTGTAATGTGATCCACTCGGGAAATGGTGTCTTACCAGGGA 3480
DB |||||
QY 3481 AAAGTGCTTACCCAGTCACGGGTCAACAGTGTGTGTGAGGATCAAAACATGGCTCTGT 3540
DB |||||
QY 3481 AAAGTGCTTACCCAGTCACGGGTCAACAGTGTGTGTGAGGATCAAAACATGGCTCTGT 3540
DB |||||
QY 3541 GAAATACTGCAACCCATTTCTATTTGGTGTGATTTCTCAGGAGTCTGAATCTTCCCTCAG 3600
DB |||||
QY 3541 GAAATACTGCAACCCATTTCTATTTGGTGTGATTTCTCAGGAGTCTGAATCTTCCCTCAG 3600
DB |||||
QY 3601 AGTCTCTTCTTCCCAACCCCTTACCAGAGCAACACTGGGATTTGAACTTCTCTGACTC 3660
DB |||||
QY 3601 AGTCTCTTCTTCCCAACCCCTTACCAGAGCAACACTGGGATTTGAACTTCTCTGACTC 3660
DB |||||
QY 3661 TCTTCTTCCCTCAGGCTGACAACTAATGTCTCTCGGGACACCCAGCTAGGGCTTCC 3720
DB |||||
QY 3661 TCTTCTTCCCTCAGGCTGACAACTAATGTCTCTCGGGACACCCAGCTAGGGCTTCC 3720
DB |||||
QY 3721 CAACCTCTTATPCAGCTGAACCTTGGATTTCTCCCAACAGGGCTTGACCTGGGTCTGTT 3780
DB |||||
QY 3721 CAACCTCTTATPCAGCTGAACCTTGGATTTCTCCCAACAGGGCTTGACCTGGGTCTGTT 3780
DB |||||
QY 3781 GGTCCCACTGGCCCAACACACATCTTTGGCCAGATTTGGGATTTCTCAATAGATTTTATAGA 3840
DB |||||
QY 3781 GGTCCCACTGGCCCAACACACATCTTTGGCCAGATTTGGGATTTCTCAATAGATTTTATAGA 3840
DB |||||
QY 3841 CATTATTCTCCACAGACTTTTAAACATGGCTTGTGTCTTCCATACACATCCGGTCTAGA 3900
DB |||||
QY 3841 CATTATTCTCCACAGACTTTTAAACATGGCTTGTGTCTTCCATACACATCCGGTCTAGA 3900
DB |||||
QY 3901 TTTAAACATTTTATATACCAAGAAATTAACCAAGCAATAGAGTACTTTTCAGATATA 3960
DB |||||
QY 3901 TTTAAACATTTTATATATACCAAGAAATTAACCAAGCAATAGAGTACTTTTCAGATATA 3960
DB |||||
QY 3961 AACTGTGTTTCTATCTTTATGATGATGCTGTGTATGATAGGGGTATGACCTGGCTGA 4020
DB |||||
QY 3961 AACTGTGTTTCTATCTTTATGATGATGCTGTGTATGATAGGGGTATGACCTGGCTGA 4020
DB |||||
QY 4021 AGTAATATTAACCATAGCTCTGGGAGGATTTACAGACCTTTTGCACTTTATGCTTTTTTG 4080
DB |||||
QY 4021 AGTAATATTAACCATAGCTCTGGGAGGATTTACAGACCTTTTGCACTTTATGCTTTTTG 4080
DB |||||
QY 4081 TGAACCTGATAACCAATGGTCAATTAATTAAGCAATATACTGGCAATTTCTGTGAATAAAC 4140
DB |||||
QY 4081 TGAACCTGATAACCAATGGTCAATTAATTAAGCAATATACTGGCAATTTCTGTGAATAAAC 4140
DB |||||
QY 4141 ATGCATATGATCTTAAAAAATAAAAAAAAAAAAAA 4175
DB |||||
QY 4141 ATGCATATGATCTTAAAAAATAAAAAAAAAAAAAA 4175
DB |||||

RESULT 2

US-10-188-832-21
; Sequence 21, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Nataasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08

```
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 4202
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-21

Query Match      17.3%; Score 723.8; DB 16; Length 4202;
Best Local Similarity 84.9%; Pred. No. 6e-200;
Matches 823; Conservative 0; Mismatches 142; Indels 4; Gaps 1;

QY 10 GCTGCTCCGATGTCAGAGCGGCATGGGCGCCCGCCACTGTGGGACCACTGCGGGC 69
Db 52 GCTGCTCCAATGCCCGAGTGGCCATGGGCGCCCGCCACTGTGGGACCACTGCGAGC 111
QY 70 TGGCAGTTCGAGGTGGATTGGTGGGAGGACAACTACACTATGCTGCTGCGCATGGCGA 129
Db 112 TGGTAGCTCGAGGTGGACTGGTGGGAGGACAACTACACCATGCTGCTGCTATCGCGA 171
QY 130 GTTCTCAACACGATCAGCAAGCTTGTGTTTCAATTTTACCTCCCATCTGCATGCTT 189
Db 172 GTTCTCAACACGATCAGCAATGCTTATTTTCAATTTTACCGCCCATCTGCTGCTT 231
QY 190 GTTCCGCGATGACGAAAGTGTTCACAGCGGCATCTACTTAATATGAGCGTCTTAGT 249
Db 232 GTTTCGTGAGTATGCAACATGCTTCAACAGTGGCATCTACTTAATCTGGACTCTTTTGT 291
QY 250 TGTAGTGGGATGGATCTGCTACTTCCATGCAAGCTGAGTTCTTGGTTCAGATGCT 309
Db 292 TGTAGTGGGATGGATCTGCTACTTCCATGCAAGCTTGTGTTTGGTTCAGATGCT 351
QY 310 TGATGAATCTGCCATCTGTGGGTTCTGATGTGTGTTGGCCATGTGGTTTCCAGGAG 369
Db 352 TGATGAATCTGCCATCTGTGGGTTCTGATGTGTGTTGGCCATGTGGTTTCCAGGAG 411
QY 370 GTATTTACCAAGATCTTTCGGAATGACAGGGCAGGTTCAAGGAGTGTGTGTCT 429
Db 412 GTATTTACCAAGATCTTTCGGAATGACAGGGGAGGTTCAAGGAGTGTGTGTCT 471
QY 430 GTCTGCAATTAACAGTGTGTTGGGTTTATCAAGCCCGCCATCAACAATATTTCCCTGAT 489
Db 472 GTCTGCGGTTACAGCGTGTGCTGCAATTTGTCAGCCCTGCCATCAACAATCTCTCTGAT 531
QY 490 GATTCTGGGACTTCCATGCACTGCGCTGCTTGTTCAGAGCTGAAGAGGTGTGACAATGT 549
Db 532 GACCTTGGGAGTTCTTTCGACTGCACTGCTCATCGCAGACTAAGAGGTGTGACAATGT 591
QY 550 GCGTGTGTTAAGTGGGCTCTTCTGCGCTCTGCTGAGTGTGCTCTCTCTCTCTCTCTGCTG 609
Db 592 GCGTGTGTTAAGTGGGCTCTTCTGCGGCTCTGCTGAGCCCTGCTGAGCCCTGCTGCTG 651
QY 610 GATCAGCGACCAAGCTTCTGTCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 669
Db 652 GATCAGTACCGAGCTTCTGCGAGTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTG 711
QY 670 TGTGTGCAATTTCTCATCTGCTTCTGTTACCTGGGCTGTGTGTGCTTCTGCTTCTT 729
Db 712 CATGTGGCACAATCTCATCTGCTTCTGCTGCTTACCTGGGCTGTGTGTGCTTCTGCTTCTT 771
QY 730 TGAATGCTGCTCAGAGTACTGAGCAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTG 789
Db 772 TGAATGCTGCTCAGAGTACTGAGCAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTG 831
QY 790 ATGGGCTTTTATTTGGTGTCTCTTATGTGCTCTTCTGTTGCTGCTGCTGCTGCTGCTGCT 849
Db 832 ATGGGCTTTTATTTGGTGTCTCTTATGTGCTCTTCTGTTGCTGCTGCTGCTGCTGCTGCT 891
QY 850 CAAGATCAGTGTATGGCAAGGAGTGTGACCAAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 909
```

```
Db 892 CAAGATCAGTGTATGGCAAGTGTGCTGCTTCTCTGCTTATGCGCCCTC-----ATGC 947
QY 910 GCTGGGCTTCTGTTGCTAGCAAGATGCTGAGGGGTTGAGGAATTTGTTGTTGTTGTTGGG 969
Db 948 AGTGGGCTTCTTGTGCTAGGAGACACCAAGGAGTTCGAATAGTTGGGTTGTTGGGCTA 1007
QY 970 TGTGTTAAAA 978
Db 1008 TCITTTCAA 1016

RESULT 3
US-10-295-027-151
; Sequence 151, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 151
; LENGTH: 4212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-151

Query Match      17.3%; Score 723.8; DB 15; Length 4212;
Best Local Similarity 84.9%; Pred. No. 6e-200;
Matches 823; Conservative 0; Mismatches 142; Indels 4; Gaps 1;

QY 10 GCTGCTCCGATGTCAGAGCGGCATGGGCGCCCGCCACTGTGGGACCACTGCGGGC 69
Db 62 GCTGCTCCAATGCCCGAGTGGCCATGGGCGCCCGCCACTGTGGGACCACTGCGAGC 121
QY 70 TGGCAGTTCGAGGTGGATTGGTGGGAGGACAACTACACTATGCTGCTGCTGCTGCTGCTG 129
Db 122 TGGTAGCTCGAGGTGGACTGGTGGGAGGACAACTACACCATGCTGCTGCTATCGCGA 181
```


[illegible]

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(825)
US-10-017-410-3

```

Query Match	15.9%;	Score 564.8;	DB 13;	Length 828;
Best Local Similarity	87.7%;	Pred. No. 3.8e-183;		
Matches 726;	Conservative 0;	Mismatches 102;	Indels 0;	Gaps 0;
Qy	35	ATGGGCGCCCGCACTGCTGGGACCACTGCGGCTGGCAGTTTCGGAGGTGGATTTGGTGC	94	
Db	1	ATGGGCGCCCGCACTGCTGGGACCACTGCGGCTGGCAGTTTCGGAGGTGGATTTGGTGC	60	
Qy	95	GAGGACAACTACACTATGCTGCTGCGCATTTGCCGAGTTCTACAAACGATCAGCAAGTGC	154	
Db	61	GAGGACAACTACACTATGCTGCTGCTATGCGCGAGTTCTACAAACGATCAGCAAGTGC	120	
Qy	155	TTGTTTTCATTTTACCTCCCATCTGCATGTGCTTGTTCGCCAGTAGCAAACTGCTTC	214	
Db	121	TTATTTTCATTTTACCGCCCATCTGCATGTGCTTGTTCGATGATGATGCAATGCTTG	180	
Qy	215	AACAGCGGCATCTACTTAAATATAGACGCTCCTAGTTGTAGTGGGATTTGATCTGTCTAC	274	
Db	181	AACAGTGACATCTACTTAAATCTGACCTCTTTTGGTTGTAGTGGGAAATTGGATCCGTCTAC	240	
Qy	275	TTCCATGCAAGCTGAGTTTCTGGGTCCAGATGCTTGATGAATCTTGCAATCTCTGGGTT	334	
Db	241	TTCCATTTTACCTTTAGTTTCTGGGTCCAGATGCTTGATGAATCTTGCAATCTCTGGGTT	300	
Qy	335	CTGATGTGTCTTTGGGCATCTGTTTTCGCCAGGAGTATTTACCAAAAGATCTTTTCGGAAT	394	
Db	301	CTGATGTGTCTTTGGCCATGTGCTTCCCCAGAGAGTATCTACCAAAGATCTTTTCGGAAT	360	
Qy	395	GACAGGGCAGGTTCAAGGCAGTGTGTGTGCTGTCTGTCTGCAATTAACAACTGCTTGCGC	454	
Db	361	GACAGGGTATAGTTCAAGGTTGGTGTGTCAGTGTCTGTCTGCGGTTTACGACGTGCTGCGCA	420	
Qy	455	TTTATCAAGCCCGCCATCAACAATATTTTCCCTGATGATTTCTGGACATTTCCATGCACTGCG	514	
Db	421	TTTGTCAAGCCTGCGCATCAACAATCTCTCTGATGACCTTGGAGTTCCTTTGACTGCA	480	
Qy	515	CTGCTTGTTCAGAGCTGAAGAGGTGTGACAAATGTGGTGTGTTTAAAGCTTGGGCTCTTTC	574	
Db	481	CTGCTCATCGCAGAGCTTAAAGAGGTGTGACAAATGCGTGTGTTTAAAGCTTGGGCTCTTTC	540	
Qy	575	TCTGSCCTCTGTGTGACTCTGGCTCTCTCTGCTGGATCAGCAACCAAGCCTCTCTGTGAG	634	
Db	541	TCGGGCTCTGTGTGGACCTTGSCCTCTGTTCTGCTTGGATCAGTGACCGAGCTTTCTGGAG	600	
Qy	635	CTGCTCTCTCTTTCATCTCCCTACCTGCACTGTGTGTGGCATAATTCATCATCTGCTTT	694	
Db	601	CTGCTGTATCTCTTCAACTTCCCTACCTGCACTGTGTGGCACAATCTCATCTGCTTT	660	
Qy	695	GCTTGTGTAACCTGGGCTGTGTGCTTTCGCTACTTTTGTATGCTGCTCAGAGATACCTGAG	754	
Db	661	GCTGCTTACCTGGGCTGTGTATGCTTTTGGCTACTTTTGTATGCTGCTCAGAGATTCCTGAG	720	
Qy	755	CAAGGTCAGTTCATCAGATTTCTGGCCCAAGGAGAAATGGGCTTTTATTGGTGTCCCTTAT	814	
Db	721	CAAGGCTGTTCATCAAGTTCTGGCCCAATGAGAAATGGGCTTTCATTGGTGTCCCTAT	780	
Qy	815	GTGTCCCTTCTGTGTGCCCAAGAGATGCGCACTCAAGATCAAGTGA	862	
Db	781	GTGTCCCTTCTGTGTGCCCAAGAGAAATCATCAGTCAAGATCAAGTGA	828	

RESULT 5
US-10-182-447-6

Db 658 GCTGCTACCTGGGCTGTTGTTATGCTTTTGCCTACTTTTGTGTCCTCAGATTCCTCTGAG 717
 Qy 755 CAAGGTCCAGTTCATCAGATTTCTGGCCCGCAGAGAAATGGGCTTTTATTGGTGTCCTCTTAT 814
 Db 718 CAAGGCCCTGTGTCATCAAAATTTCTGGCCCGCAGAGAAATGGGCTTTCATTTGTGTCCTCTAT 777
 Qy 815 GTGTCCCTTCTGTGTGTCGCCCAAGAAAGTGGCAGTCAAGATCAAG 859
 Db 778 GTGTCCCTCTGTGTGTCGCCCAAGAAATTCATCAGTCAAGACCAG 822

 RESULT 6
 US-09-917-800A-119/C
 ; Sequence 119, Application US/09917800A
 ; Patent No. US20020119462A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark
 ; APPLICANT: Johnson, Kory
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5038-US
 ; CURRENT APPLICATION NUMBER: US/09/917.800A
 ; CURRENT FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,040
 ; PRIOR FILING DATE: 2000-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,880
 ; PRIOR FILING DATE: 2000-11-02
 ; PRIOR APPLICATION NUMBER: US 60/290,029
 ; PRIOR FILING DATE: 2001-05-11
 ; PRIOR APPLICATION NUMBER: US 60/290,645
 ; PRIOR FILING DATE: 2001-05-15
 ; PRIOR APPLICATION NUMBER: US 60/292,336
 ; PRIOR FILING DATE: 2001-05-22
 ; PRIOR APPLICATION NUMBER: US 60/295,798
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 60/297,457
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,884
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: US 60/303,459
 ; PRIOR FILING DATE: 2001-07-09
 ; NUMBER OF SEQ ID NOS: 1740
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 119
 ; LENGTH: 595
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA848826
 US-09-917-800A-119

 Query Match 9.6%; Score 400.4; DB 9; Length 595;
 Best Local Similarity 87.9%; Pred. No. 8.3e-106;
 Matches 531; Conservative 0; Mismatches 56; Indels 17; Gaps 8;

 Qy 3557 ATTCTATTGGTTGGGATTTCTCAGGAGTCTGAATCTTCCCTCAGAGTCTCTTCTCCCA 3616
 Db 595 ATTCTATTGGTTGGACTTCTCAGGTGTCTGAAATCTCCCTCAGAAATCTCTTCTCTTAA 536
 Qy 3617 CCCCTACCCAGAGCCAAACATCGGATTGGAACCTTC-----TCTGACTCTCTTCTTCCCTC 3672
 Db 535 CCGCCTCCCAATC--CACTGGGATTTGAACCTTCAAATCCCTGCCTCTCTTCTTCCCTTC 479
 Qy 3673 AGGTCTCACAACTAATGCTCTCTGGGGACACCCAGCTAGGCGCTTCCCCAACTCCCTT-AT 3731
 Db 478 AGGTCTGACCACTAAACGGTCTCTGGGGACACCTAACTAGGCGCTTCCCCAGCTCTCTTCT 419
 Qy 3732 CCAGCTGAACCTTGATTTCTTCCCAACACAGGGCTTGACCTGGGTGCTGTTGGTCCCACTGG 3791
 Db 418 CCAGCCAAAATTGGATTCTTCCCAACACAGGGCTTGACCTGGGTGCTGTTGGTCCCACTGC 359

QY 3701 -----CACCCAGCTAGGG--CCTTCCCAACTCTTATCC-----AGCTGA 3739
DB 652 TGGATGCTGAAGCTCACCAGTCAGGCGCCCTCTCCTAGTCTCTTTTACACTGAAATTA 593
QY 3740 ACTTGGATTCTTCCCAACCGGCTTGACCTGGGTGCTGTGGTCCCACTGCGCAACAC 3799
DB 592 TCTGAAAGCTTTTATAGCAAGGCTTTGCGCTAGGTGCTATTATTTCCAGCTGCGCAAGAG 533
QY 3800 ACATCTTTGGCCAGATTGGGATTTCTCAATAGATTTTATAGACATTAATTTCTCCACAGACT 3859
DB 532 AAGTCTTTGGCCAGATTGGGATTTCTCAATAGATTTTATAGACATTAATTTCCCTGCAACT 473
QY 3860 TTAACAAC-----TGGCTTGTGCTTTCCATAC 3887
DB 472 TAAAAAATAAATAACCCCTACTTTATAGGACTAATTTGTTGAATTTGATCTTTCTCTGT 413
QY 3888 ACATCCGGTCAGATTTAAACTATTTTATAACCCAGGAAATTAACCAAGCAAAATAGAGT 3947
DB 412 ATGTAAACCCAGATTTAAACTATTTTATAACCACAATATGTAATCAGAGCA-ATATAGT 354
QY 3948 ACTTTTCAGATATAAACTGTGTTTCATCTTTATGTAGAGTGTGCTATGTATAGGCGGTAT 4007
DB 353 GTTTTCAGATATAACCTGTTTATATACCTTTATGTAG-GTGTCTACATAAGGGTGGCAT 295
QY 4008 GTACCTCGGTGAAGTAAATAAAC-----CATAGCTCTGGAGGAGATTTACAGAC-----4057
DB 294 GCCCACTGGCTGTGGTAAATTTAACTCTCAATGCTTTGGGAGTGACTTAAAGGCTTTTG 235
QY 4058 -----CTTTTGCACATTTATGCTTTT-TGTGAACCTCTGATACCATGCTCAATATTA 4108
DB 234 AAGTGGAGCTTTTGACATTTATACCTTTTCTGTGAACATGATGATTAATTTGATATTA 175
QY 4109 AAGCAATAAATCGGCAATTTCTGTGAATAAACAATGATATGTATCT 4154
DB 174 AAGCTG-TAAGTGGCAATTTTACGCAATGAATATGATATGTTGT 130

RESULT 11

US-10-106-698-611
; Sequence 611, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 611
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-611

Query Match 3.0%; Score 125.8; DB 15; Length 1473;
Best Local Similarity 62.4%; Pred. No. 4e-25;
Matches 440; Conservative 0; Mismatches 177; Indels 88; Gaps 12;

QY 3527 AAACATGGCTCTGTGAAATATCTGCCACCATCTATTTGGTGGATTTCTCAGGAGTCTG 3586
DB 637 ACACAGGGCTTTGTAAATAATACGACCACTATTTCCACTTACTGGATCTGTCAGGTGTGA 696
QY 3587 AATCTTCCCTCAGAGTCTCTTCTCCCAACCCCTACCAGAGCCCAACACTGGGATTTGA 3646
DB 697 AAACCTCTCTGCCAG-----TTTCATATGCTTCCATGAGCCCTCAGGACTGGGATTTGA 751

QY 3647 ACCTTCTCTGACTCTCTTCTTCCCTCAGGCTCTGACAACTAATGCTCTCTGGGA-----3700
DB 752 GCCTTCTCGG-----CTCTTTATCCCTTGGGCGAGACATGGAACCAATCTCTGAGGACACAGG 808
QY 3701 -----CACCCAGCTAGGG--CCTTCCCAACTCTTATCC-----AGCTGA 3740
DB 809 TGGATGCTGAAGCTCACCAGTCAGGCGCCCTCTCCTAGTCTCTTTTACACTGAAATTA 868
QY 3741 CTTCGATTTCTTCCCAACCGGCTTGACCTGGGTGCTGTGGTCCCACTGCGCAACACAC 3800
DB 869 TCTGAAAGCTTTTATAGCAAGGCTTTGCGCTAGGTGCTATTATTTCCAGCTGCGCAAGAG 928
QY 3801 CATCTTTGGCCAGATTTGGGATTTCTCAATAGATTTTATAGACATTAATTTCTCCACAGACT 3860
DB 929 AGTCTTTGGGCGAGATTTGGGATTTCTCAATAGATTTTATAGACATTAATTTCCCTGCAACT 988
QY 3861 TAAACA-----TGGCTTGTGCTTTCCATACA 3888
DB 989 AAAAAAATAAATAACCCCTACTTTATAGGACTAATTTGTTGAATTTGATCTTTCTCTGTA 1048
QY 3889 CATCCGGTCAGATTTAAACTATTTTATAACCAAGGAAATTAACCAAGCAAAATAGAGT 3948
DB 1049 TGTAAACCCAGATTTAAACTATTTTATAACCACAATATGTAATCAGAGCA-ATATAGT 1107
QY 3949 CTTCAGATATAAACTGTGTTTCATCTTTATGTAGAGTGTGCTATGTATAGGCGGTATG 4008
DB 1108 TTTTCAGATATAATACCTTTGTTTATATACCTTTATGTAG-GTGTCTACATAAGGGTGGCAT 1166
QY 4009 TACCTCGCTCAAGTAATATTAAC-----CATAGCTCTCGGAGGATTTACAGAC-----4057
DB 1167 CCCACTGGCTGTGTAAATTTAATCCTCATTTGCTTTGGAGTGACTTAAAGGCTTTTGA 1226
QY 4058 -----CTTTTGCACATTTATGCTTTT-TGTGAACCTCTGATACCATGCTCAATATTA 4109
DB 1227 AGTGGAGCTTTTGCACATTTATACCTTTTCTGTGAACATGATGATATTTGATATTA 1286
QY 4110 AGCCAATAAATCGGCAATTTTCTGTGAATAAACAATGATATGTATCT 4154
DB 1287 AGCTG-TAAGTGGCAATTTTACGCAATGAATATGATATGTTGT 1330

RESULT 12

US-10-242-535A-50197
; Sequence 50197, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50197
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-50197

Query Match 1.6%; Score 65; DB 16; Length 418;
Best Local Similarity 55.4%; Pred. No. 1.1e-07;
Matches 201; Conservative 0; Mismatches 140; Indels 22; Gaps 3;

QY 3527 AAACATGGCTCTGTGAAATACTGCCACCACTTCTATTGTTGGATTTCTCAGGAGTCTG 3586

Db 27 ACACAGGCTTTGTAATAATACGACCACTATTTCACCTTACTGGATCTGTCAAGGTGTA 86
QY 3587 AATCTTCCCTCAGGAGTCTCTTCTCCCAACCCCTACCCAGAGCCCAACACTGGGATTTGA 3646
Db 87 AAACCTTCTCGCCAGTTTCATCATGCTTCCATGAGCCCTCAGGACTGGGATTTGAGCCTT 146
QY 3647 ACCTTCTCTGACTCTCTTCTTCCCTCAGGTCTGACAACTAAT-----GGTCTCTGG 3697
Db 147 CTGGCTCTTTATCCCTTGGGGCAGGACATGGAACCATCTCTGAGGGACCAAGTGGATGC 206
QY 3698 GGACACCCAGCTAGGCGCTTCCCAACTCCTTATCCAGCTGAACCTTGGATTTCTCCCAA- 3756
Db 207 TGAAGCTCACCCAGTCAAGGGCCCTCTCCTAGCTCTCTTTACACTGAAATTAATCTGAA 266
QY 3757 -----CCAGGGCTTCCACCTGGGTGCTGTGGTCCACCTGGCCCAACACACATCT 3805
Db 267 GCTTTCATAGCCAGGCTTCCCTAGGGTGTCTATTATCCAGCTGGCCAAAGAGA-AGCT 325
QY 3806 TTGGCCAGATTGGGATTTCTAATAGATTTTATAGACATTTATCTCCACAGACTTTAAAA 3865
Db 326 TGGGCCAGATTGGGATTTCTAATGATTTTATAGACATTTATCTCCCTGCAAACTTTAAAA 385
QY 3866 CAT 3868
Db 386 AAT 388

RESULT 13

US-10-085-783A-50197
; Sequence 50197, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50197
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-50197

Query Match 1.6%; Score 65; DB 16; Length 418;
Best Local Similarity 55.4%; Pred. No. 1.1e-07;
Matches 201; Conservative 0; Mismatches 140; Indels 22; Gaps 3;

QY 3527 AACATGGCTCTGTGAAATACGCCACCATTTATTTGGTGGATTTCTCAGGATCTG 3586
Db 27 ACACAGGCTTTGTAATAATACGACCACTATTTCACCTTACTGGATCTGTCAAGGTGTA 86
QY 3587 AATCTTCCCTCAGGAGTCTCTTCTCCCAACCCCTACCCAGAGCCCAACACTGGGATTTGA 3646
Db 87 AAACCTTCTCTCGCCAGTTTCATCATGCTTCCATGAGCCCTCAGGACTGGGATTTGAGCCTT 146
QY 3647 ACCTTCTCTGACTCTCTTCTTCCCTCAGGCTGACAACTAAT-----GGTCTCTGG 3697
Db 147 CTGGCTCTTTATCCCTTGGGGCAGGACATGGAACCATCTCTGAGGGACCAAGTGGATGC 206
QY 3698 GGACACCCAGCTAGGCGCTTCCCAACTCCTTATCCAGCTGAACCTTGGATTTCTCCCAA- 3756
Db 207 TGAAGCTCACCCAGTCAAGGGCCCTCTCTAGCTCTCTTTACACTGAAATTAATCTGAA 266

QY 3757 -----CCAGGGCTTGCCTGGGTGCTGTGGTCCCACTGGCCCAACACACATCT 3805
Db 267 GCTTTCATAGCCAGGCTTTCCTAGGGTGTCTATTATTCAGCTGGCCAAAGAGA-AGCT 325
QY 3806 TTGGCCAGATTGGGATTTCTAATAGATTTTATAGACATTTATCTCCACAGACTTTAAAA 3865
Db 326 TGGGCCAGATTGGGATTTCTAATGATTTTATAGACATTTATCTCCCTGCAAACTTTAAAA 385
QY 3866 CAT 3868
Db 386 AAT 388

RESULT 14

US-10-322-281-265/c
; Sequence 265, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 265
; LENGTH: 77530
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(77530)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-265

Query Match 1.5%; Score 63.6; DB 17; Length 77530;
Best Local Similarity 63.2%; Pred. No. 9.2e-06;
Matches 115; Conservative 0; Mismatches 64; Indels 3; Gaps 1;
QY 2701 GGGGCTGGAGAGTTGGTTCAACAGTTAAGAACCCGCTGTCTCTCTCCACAGACGCCAG 2760
Db 30410 GAGGCTGGAGAGAGGCTCATCATGTTAGGAGCACTGGCTCTCTCTGAGGGCCCAACA 30351
QY 2761 CGACTTGTGAGGCTCATAACCA---GCAACTCCAGCCCTGGGGATCTGACGCCCTCTTC 2817
Db 30350 CTCACATGGTGGCTCACACACCACTCCCACTCCCAAGAGATATGATGCCCTCTTC 30291
QY 2818 TGACTTCTAAGGGCAATTTGGTGCACACGTACATACAGGCGAGGCAAAACATTATACATGT 2877
Db 30290 TAGCCTCCGACACACACAGGGGTACATACATACAGCAAGCAGAACATTCATACAT 30231
QY 2878 AA 2879
Db 30230 AA 30229

RESULT 15

US-09-771-208-20/c
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFORD, ERIC
; APPLICANT: HORVAT, SIMON
; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407T-923710US
; CURRENT APPLICATION NUMBER: US/09/771,208
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; PRIOR FILING DATE: 1997-12-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 20
; LENGTH: 659158
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (546998)..(547017)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (494715)..(494814)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (390986)..(391005)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (346860)..(346883)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (317174)..(317193)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (280353)..(280373)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (271829)..(271848)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (183872)..(183891)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (170625)..(170645)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (132680)..(132700)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc feature
; OTHER INFORMATION: n is a, c, g, or t
; US-09-771-208-20

Query Match      1.5%; Score 61.4; DB 9; Length 659158;
Best Local Similarity 60.9%; Pred. No. 0.00017;
Matches 137; Conservative 0; Mismatches 81; Indels 7; Gaps 2;

Qy 2699 CTGGGCTGGAGAGTTGTTCAACAGTTAAGAACCCGTGTGCTCTCTCCAGACGACCC 2758
Db 318008 CTAGGGCTGGTGAGATGGCTCAGCAGTTAAGAGCACTTGCTGCTCTCTAAGAACCCAC 317949

Qy 2759 AGCGACTTGTGAGGCTCATACACGCA---ACTCCAGCCCTGGGGCATCTGACGCCCTC 2814
Db 317948 ACCTACCTATGTGGTGGCTAATACAGACACTCCGGTCCAGGGGATGTGATGCCCTC 317889

Qy 2815 TTCTGACTTCTAAGGGCATTGGTGCACA---CGTACATACAGGAGGCAAAACATTTAT 2871
Db 317888 TTCTGGCTCTCGAGGGCACTGCATACACATGTTGCACAGGTATGCAGAGGAATATCCAC 317829

Qy 2872 ACATCTAAGTAATAAATGCAATAGTTAGTGACGGCTGAAGGA 2916
Db 317828 ACACATAAAATTTAAATCAAAATATTTCTTAGGAGGGGAGAGA 317784
```

Search completed: November 21, 2004, 07:43:59
Job time : 1957.22 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:50:47 ; Search time 11519.4 Seconds
(without alignments)
13206.895 Million cell updates/sec

Title: US-10-017-410-1
Perfect score: 4175
Sequence: 1 ggcacagggtctcgat.....aaaaaaaaaaaaaac 4175

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3028.8	72.5	3071	3 AK035314	AK035314 Mus muscu
2	1341	32.1	1341	3 BC023423	BC023423 Mus muscu
3	780	18.7	1022	5 BU511164	BU511164 AGENCOURT
4	714.4	17.1	717	7 CN537424	CN537424 UI-M-HS0-
5	681.2	16.3	688	2 BB108761	BB108761 BB108761
6	681	16.3	681	9 AY401891	AY401891 Mus muscu
7	670.6	16.1	1388	6 CF110927	CF110927 Shultzoni
8	668.4	16.0	670	7 CN721804	CN721804 E0830A03-
9	667.6	16.0	1173	3 AK085306	AK085306 Mus muscu
10	662.6	15.9	955	1 AV227941	AV227941 AV227941
11	653.6	15.7	868	6 CA976684	CA976684 AGENCOURT
12	631.4	15.1	797	6 CA463294	CA463294 AGENCOURT
13	616.2	14.8	651	2 BB627599	BB627599 BB627599
14	594	14.2	617	5 BM942111	BM942111 UI-M-CG0p
15	576.8	13.8	606	8 A2411514	A2411514 IM0184807
16	575.2	13.8	1124	4 BI080821	BI080821 602878593
17	568.2	13.6	916	2 BF144307	BF144307 601787125
18	563.6	13.5	633	2 BB660847	BB660847 BB660847
19	559.4	13.4	562	2 AW701866	AW701866 uq95g09.x
20	552	13.2	552	2 AW490555	AW490555 UI-M-BH3-
21	548	13.1	548	5 BX523075	BX523075 BX523075
22	543.4	13.0	681	9 AY401890	AY401890 Pan trogl
23	541.8	13.0	681	9 AY401889	AY401889 Homo sapi
24	497.2	11.9	731	6 CA503516	CA503516 UI-R-F90-

25	497	11.9	565	6	CF115220	Shultzoni
26	480	11.5	480	2	BF662488	maab6c03.
27	479.6	11.5	1072	5	BQ715126	AGENCOURT
28	476.6	11.4	548	6	CF169369	B0812G07-
29	475	11.4	492	1	AA212916	mw86f03.r
30	474	11.4	490	1	AA212925	mw86h04.r
31	471.2	11.3	528	2	BB764116	BB764116
32	470.8	11.3	485	1	AA475320	vh14a03.r
33	470	11.3	505	2	BB752973	BB752973
34	465	11.1	690	4	BM387041	UI-R-CN1-
35	440	10.5	440	2	BF661433	maab6c03.
36	437.8	10.5	442	2	BF468547	UI-M-BH3-
37	431.8	10.3	448	1	AA462019	v95sd02.r
38	425.6	10.2	432	6	CF169808	CF169808
39	419.2	10.0	622	6	CB723138	UI-M-GH0-
40	417.2	10.0	422	4	BG230006	mac25g04.
41	412.2	9.9	769	6	CA510569	UI-R-FJ0-
42	410	9.8	427	2	AW702104	uq95g09.y
43	408	9.8	408	1	AA462654	AA462654
44	405	9.7	498	4	BI848265	470659 MA
45	400.4	9.6	595	1	AA848826	EST191587

ALIGNMENTS

RESULT 1	AK035314	3071 bp	linear	HTC 03-APR-2004
LOCUS	Mus musculus adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:9530014B21 product:similar to CANCER RELATED GENE-LIVER 1 [Mus musculus], full insert sequence.			
DEFINITION	AK035314			
ACCESSION	AK035314			
VERSION	AK035314.1	GI:26084521		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Mech. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE 2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE 3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
MEDLINE	20530913			
PUBMED	11076861			
REFERENCE 4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
MEDLINE				
PUBMED				
REFERENCE 5				
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research			

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
of full-length cDNAs

Nature 420, 563-573 (2002)
6 (bases 1 to 3071)

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashida, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komori, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

source
1. 3071
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9530014B21"
/db_xref="taxon:10090"
/clone="9530014B21"
/sex="male"
/tissue_type="urinary bladder"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1. 3071
/notes="similar to CANCER RELATED GENE-LIVER 1 [Mus
musculus] (SPTRIAL40408, evidence: FASTA, 100%ID,
77.8%length, match=642)";

ORIGIN

Query Match 72.5%; Score 3028.8; DB 3; Length 3071;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3065; Conservative 0; Mismatches 2; Indels 5; Gaps 3;

QY 1065 AGTATCTTCCAGGCATAGGGCTGCTGGGTTATGTCATATGACATCAACAGTA 1124
DB 2 AGTATCTTCCAGGCATAGGGCTGCTGGGTTATGTCATATGACATCAACAGTA 61
QY 1125 GTTCTTTGGAGACCTAGGGCAACCCAAAGTTTCTGCTGGAGAGGGTAGTTCGTGTT 1184
DB 62 GTTCTTTGGAGACCTAGGGCAACCCAAAGTTTCTGCTGGAGAGGGTAGTTCGTGTT 121
QY 1185 CCACCATTTCCAGACTCTCAACCCCTGATAACACAGACCTTCTGATTTGGTGATGAAG 1244
DB 122 CCACCATTTCCAGACTCTCAACCCCTGATAACACAGACCTTCTGATTTGGTGATGAAG 181
QY 1245 GTTCCAGAACTTTTCATTTTCCGGGAGAACTGCTCTCAACAAACCAAGTGGGCAAA 1304
DB 182 GTTCCAGAACTTTTCATTTTCCGGGAGAACTGCTCTCAACAAACCAAGTGGGCAAA 241
QY 1305 ACACGTGTGGGGGTGTGATAAGACGGGTGCTGTCTACCTCGCACTCTTACCTTCCC 1364

DB 242 ACACGTGTGGGGGTGTGATAAGACGGGTGCTGTCTACCTCGCACTCTTAGCTTTCCC 301
QY 1365 AGGTTCCTCCACGTTGTTTGGATCCTTCCACTACCTCTGCTGAGAGATGAGCCACGGCT 1424
DB 302 AGGTTCCTCCACGTTGTTTGGATCCTTCCACTACCTCTGCTGAGAGATGAGCCACGGCT 361
QY 1425 TCAGAGGGCAAGCTGGCAACACCTCTATGCAAAAGCTACACTCTCTTTTAAGACAC 1484
DB 362 TCAGAGGGCAAGCTGGCAACACCTCTATGCAAAAGCTACACTCTCTTTTAAGACAC 421
QY 1485 ATTAC--ATAAGACACTATTCTCTCTCCAGAGTGCAGAGCTCAGACCCACAGAGAA 1542
DB 422 ATTCATATAGACACTATTCTCTCTTCAGAGTGCAGAGCTCAGACCCACAGAGAA 481
QY 1543 TCCTTCAGGTATGTGAAGATTCCACACAGCCCCCTCTTTGTGACCTCTGTGAGAGCAA 1602
DB 482 TCCTTCAGGTATGTGAAGATTCCACACAGCCCCCTCTTTGTGACCTCTGTGAGAGCAA 541
QY 1603 GGCCTCG--TTGTAATCGCGAGGACGCTCTGGAGGCTTGACTGTGGAGCGCTGTGAGA 1661
DB 542 GGCCTCGCTTGTAAATCGCGAGGACGCTCTGGAGGCTTGACTGTGGAGCGCTGTGAGA 601
QY 1662 TTTTGGAAAGCTTTTCTTATTGGAAGCTTTTGGTGTGCTTACCTCCAAAGATTCTGACC 1721
DB 602 TTTTGGAAAGCTTTTCTTATTGGAAGCTTTTGGTGTGCTTACCTCCAAAGATTCTGACC 661
QY 1722 CGCTTCTCTGCTGTTTTTAGGGGATATGTGCTTCCTGAGTGCAGTAGCTCCCTAGGAT 1781
DB 662 CGCTTCTCTGCTGTTTTTAGGGGATATGTGCTTCTGAGTGCAGTAGCTCCCTAGGAT 721
QY 1782 GTGGGCTCGGCTTTTGTTCATTAATACCTGGTGTCTTACCTGCTTCTCAGAGCACTTT 1841
DB 722 GTGGGCTCGGCTTTTGTTCATTAATACCTGGTGTCTTACCTGCTTCTCAGAGCACTTT 781
QY 1842 GCTCTTCTTGATGCTGGGGGTCACTACACTCTGATTTGGCTGGGTGGGATCCAGGGAG 1901
DB 782 GCTCTTCTTGATGCTGGGGGTCACTACACTCTGATTTGGCTGGGTGGGATCCAGGGAG 841
QY 1902 GAAGGGGAGACTCTGACATGAATGTCTCTACCTGATCTCTACTGTCTTCTACTGGGCTC 1961
DB 842 GAAGGGGAGACTCTGACATGAATGTCTCTACCTGATCTCTACTGTCTTCTACTGGGCTC 899
QY 1962 TCTTCAGGTATGAAGTGGGTATGGGTATCTCAGGGATGTTGTAACTCAGGACCTTC 2021
DB 900 TCTTCAGGTATGAAGTGGGTATGGGTATCTCAGGGATGTTGTAACTCAGGACCTTC 959
QY 2022 TGCTTCTGACATTCATTTCTCAGTGTGAGACTGCACCTCAGGGGACTGGAGTTGAAACA 2081
DB 960 TGCTTCTGACATTCATTTCTCAGTGTGAGACTGCACCTCAGGGGACTGGAGTTGAAACA 1019
QY 2082 CGTTCCTGAGTGTCTGAGTCTGAAGAGAGTACTGCGGGTCTGCGACCTTCTAGGA 2141
DB 1020 CGTTCCTGAGTGTCTGAGTCTGAAGAGAGTACTGCGGGTCTGCGACCTTCTAGGA 1079
QY 2142 TCTCTTACTCTGTTTGAACCTTTCACAGGTACAAAGTGGGACTGACCTTAAAGAGTTT 2201
DB 1080 TCTCTTACTCTGTTTGAACCTTTCACAGGTACAAAGTGGGAACTGACCTTAAAGAGTTT 1139
QY 2202 TTAATGAATAGACTTTTCTTCTGCTTTTGGGCTTTTGAAGTTTAAAGTAACTGCTGG 2261
DB 1140 TTAATGAATAGACTTTTCTTCTGCTTTTGGGCTTTTGAAGTTTAAAGTAACTGCTGG 1199
QY 2262 GCTATTTCTGTTTATAGTCAAAATATTATAGAACATGAAGGTGTAAATTAAGTTGTC 2321
DB 1200 GCTATTTCTGTTTATAGTCAAAATATTATAGAACATGAAGGTGTAAATTAAGTTGTC 1259
QY 2322 TTTTATTAATTCATAGCAATTTACCAACTCTCCAGGTAGCAAAACACACAGTCAATTGAT 2381
DB 1260 TTTTATTAATTCATAGCAATTTACCAACTCTCCAGGTAGCAAAACACACAGTCAATTGAT 1319
QY 2382 TGGCGATTTACAGGCGAGAAATCTTTGGTTTGAATGATGATGATGATGATGATGATGAT 2441
DB 1320 TGGCGATTTACAGGCGAGAAATCTTTGGTTTGAATGATGATGATGATGATGATGATGAT 1379

Qy	2442	GCTTAAATATGTTTAAATGACCTTGGAGTGAATGTTTGGGCAACCTAGGGGTTTATG	2501
Db	1380	GCTTAAATATGTTTAAATGACCTTGGAGTGAATGTTTGGGCAACCTAGGGGTTTATG	1439
Qy	2502	GATCAGAAATGGCGCTGAGTCCCTTGGTCTTGGAGTCTAGAGTTTTCAGAGGCAAAAT	2561
Db	1440	GATCAGAAATGGCGCTGAGTCCCTTGGTCTTGGAGTCTAGAGTTTTCAGAGGCAAAAT	1499
Qy	2562	CAAACTAGCCCTTGTCTGCTTCACTTACAGGCTCGGAGCTGCCCTTATATATCTA	2621
Db	1500	CAAACTAGCCCTTGTCTGCTTCACTTACAGGCTCGGAGCTGCCCTTATATATCTA	1559
Qy	2622	CATAAGGACTTATATACATAAGTCTGTATTAATGTCTGTAAGATGACCTAGCTGCT	2681
Db	1560	CATAAGGACTTATATACATAAGTCTGTATTAATGTCTGTAAGATGACCTAGCTGCT	1619
Qy	2682	TCATCTGGAAGGTCGTCGCGGCTGAGAGTTGGTTCAACAGTTAAGAACCCGCTGTTC	2741
Db	1620	TCATCTGGAAGGTCGTCGCGGCTGAGAGTTGGTTCAACAGTTAAGAACCCGCTGTTC	1679
Qy	2742	TCTCTCCAGACGACCCAGGACCTTGTGAGGCTCATACAGCAACTCCAGCCCTGGGGC	2801
Db	1680	TCTCTCCAGACGACCCAGGACCTTGTGAGGCTCATACAGCAACTCCAGCCCTGGGGC	1739
Qy	2802	ATCTGAGCCCTCTTCTGACTTCTAAGGGCATTTGGTGACACGTACATACAGCAGGCA	2861
Db	1740	ATCTGAGCCCTCTTCTGACTTCTAAGGGCATTTGGTGACACGTACATACAGCAGGCA	1799
Qy	2862	AAACATTTATACATGTAACGTAAATAATGCATAGTTAGTGAGACGCTGAAGAAAGGA	2921
Db	1800	AAACATTTATACATGTAACGTAAATAATGCATAGTTAGTGAGACGCTGAAGAAAGGA	1859
Qy	2922	GTTTATAGTCAAGGTTTAGTCTGACCTGAGTGCCCTTCCCTTAGCCGCGATGGTCCCT	2981
Db	1860	GTTTATAGTCAAGGTTTAGTCTGACCTGAGTGCCCTTCCCTTAGCCGCGATGGTCCCT	1919
Qy	2982	AAAGTCAAGGAAAGCAATTTCTTTTATACAGAAAGGACTTATCCCTAGGGCCCTCT	3041
Db	1920	AAAGTCAAGGAAAGCAATTTCTTTTATACAGAAAGGACTTATTTCCCTAGGGCCCTCT	1979
Qy	3042	GCTGACATCCCTAGGAACAGAGATAAATACGATGATGTAATGAACATGCTTGGGTG	3101
Db	1980	GCTGACATCCCTAGGAACAGAGATAAATACGATGATGTAATGAACATGCTTGGGTG	2039
Qy	3102	AAAGGAGCCGAGTACCTGACCTGGACCCAGTGGGCCACTTCCACAAGCGAAGCCCGTAAC	3161
Db	2040	AAAGGAGCCGAGTACCTGACCTGGACCCAGTGGGCCACTTCCACAAGCGAAGCCCGTAAC	2099
Qy	3162	CGATGTCACCTCAGAAATCTTCTCCCTTCTGGTAGAGATGACATCGATGATGTCGT	3221
Db	2100	CGATGTCACCTCAGAAATCTTCTCCCTTCTGGTAGAGATGACATCGATGATGTCGT	2159
Qy	3222	GACCACTGGGCGCTGATGGGTGTCGGCTCTGTTCTGTTCCGAATCTACCTGAGATCTC	3281
Db	2160	GACCACTGGGCGCTGATGGGTGTCGGCTCTGTTCTGTTCCGAATCTACCTGAGATCTC	2219
Qy	3282	AGGACAGAGGAAGCCATGAATGTTTACCAAGTGGTCATGGCTGTGATGATTTTACAGTTT	3341
Db	2220	AGGACAGAGGAAGCCATGAATGTTTACCAAGTGGTCATGGCTGTGATGATTTTACAGTTT	2279
Qy	3342	TGAACCAATATGTTTATAGGAAATTTCTCTCTCTAGTGCCCTGTGATGCAAAAGC	3401
Db	2280	TGAACCAATATGTTTATAGGAAATTTCTCTCTCTAGTGCCCTGTGATGCAAAAGC	2339
Qy	3402	CAGCCCTTCAAGTGTTCCTTCTGCTGCTTCCCTTGTATGTAATGTGATGCACTGGGAA	3461
Db	2340	CAGCCCTTCAAGTGTTCCTTCTGCTGCTTCCCTTGTATGTAATGTGATGCACTGGGAA	2399
Qy	3462	ATGGGTGTCTACCCAGGGAAGGTCCTTACCCAGTCAAGGCTCAACAGTGTGTGTGTA	3521
Db	2400	ATGGGTGTCTACCCAGGGAAGGTCCTTACCCAGTCAAGGCTCAACAGTGTGTGTGTA	2459

Qy	3522	GGATCAACATGCTCTGTGAAATATCTGCCATCTTCTATTGTTGGATTTCTCAGGA	3581
Db	2460	GGATCAACATGCTCTGTGAAATATCTGCCATCTTCTATTGTTGGATTTCTCAGGA	2519
Qy	3582	GTCTGAATCTTCCCTCAGAGTCTCTTCTTCCCAACCCCTACCCAGAGCCAACTGGGA	3641
Db	2520	GTCTGAATCTTCCCTCAGAGTCTCTTCTTCCCAACCCCTACCCAGAGCCAACTGGGA	2579
Qy	3642	TTTGAACCTTCTGACTCTCTTCTTCCCTCAGGTCTGACAACTAATGTTCTCTGGGAC	3701
Db	2580	TTTGAACCTTCTGACTCTCTTCTTCCCTCAGGTCTGACAACTAATGTTCTCTGGGAC	2639
Qy	3702	ACCAGCTAGGCTTCCCAACTCTTATCCAGCTGAACTTGATTTCTTCCCAACGAG	3761
Db	2640	ACCAGCTAGGCTTCCCAACTCTTATCCAGCTGAACTTGATTTCTTCCCAACGAG	2699
Qy	3762	GCTTGACCTGGGTCTGTTGGTCCCACTGGCCCAACACACATCTTGGCCAGATTTGGAT	3821
Db	2700	GCTTGACCTGGGTCTGTTGGTCCCACTGGCCCAACACACATCTTGGCCAGATTTGGAT	2759
Qy	3822	TCTCAATAGATTTTATAGACATTTCTCCACAGACTTTTAAACATGCTTGTCTTT	3881
Db	2760	TCTCAATAGATTTTATAGACATTTCTCCACAGACTTTTAAACATGCTTGTCTTT	2819
Qy	3882	CCATACATCCGGTCTGATTTAAACATTTTATTAACCAACAGAAATTAACCAAGCAA	3941
Db	2820	CCATACATCCGGTCTGATTTAAACATTTTATTAACCAACAGAAATTAACCAAGCAA	2879
Qy	3942	TAGAGTACTTTTATAGATATAAATCTGTTTCTATCTTATGATGAGTGTCTATGATAGG	4001
Db	2880	TAGAGTACTTTTATAGATATAAATCTGTTTCTATCTTATGATGAGTGTCTATGATAGG	2939
Qy	4002	CGGTATGATCCCTGGCTGAAATATAATTAACCATAGCTCTGGGAGGATTTACAGACCTT	4061
Db	2940	CGGTATGATCCCTGGCTGAAATATAATTAACCATAGCTCTGGGAGGATTTACAGACCTT	2999
Qy	4062	TGCACTTTATGCTTTTGTGAACTCTGATAACCATGCTCAATATTAAGCAATTAATG	4121
Db	3000	TGCACTTTATGCTTTTGTGAACTCTGATAACCATGCTCAATATTAAGCAATTAATG	3059
Qy	4122	GCATTTTCTGTG 4133	
Db	3060	GCATTTTCTGTG 3071	

RESULT 2	BC023423	1341 bp	mRNA	linear	HTC 20-SEP-2002
LOCUS	Mus musculus, clone IMAGE:5010167, mRNA.				
DEFINITION	Mus musculus, clone IMAGE:5010167, mRNA.				
ACCESSION	BC023423				
VERSION	BC023423.1	GI:19484055			
KEYWORDS	HTC.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1 (bases 1 to 1341)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-FEB-2002) National Institutes of Health, Mammalian Genome Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (ILNL) Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/				

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM14067 row: j column: 05

High quality sequence stop: 681.

FEATURES

Location/Qualifiers
1..1022
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6505924"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_134"
/notes="vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 18.7%; Score 780; DB 5; Length 1022;
Best Local Similarity 96.8%; Pred. No. 7.9e-204;
Matches 871; Conservative 0; Mismatches 20; Indels 9; Gaps 7;
QY 10 GCTGCTCCGATGCTCCAGAGCGGCATGGCGCCCGCAGCTGGTGGGACACCTGCGGGC 69
DB 95 GCTGCTCCGATGCTCCAGAGCGGCATGGCGCCCGCAGCTGGTGGGACACCTGCGGGC 154
QY 70 TGGCAGTTCGGAGTGGATTGGTGGGAGGACAACTACACTATCGTGGCTGCGCATTCGCCA 129
DB 155 TGGCAGTTCGGAGTGGATTGGTGGGAGGACAACTACACTATCGTGGCTGCGCATTCGCCA 214
QY 130 GTTCTACACAGCATGACGAACTGCTGTTTTCATTTTACCTCCCATCTGCATGTCGCTT 189
DB 215 GTTCTACACAGCATGACGAACTGCTGTTTTCATTTTACCTCCCATCTGCATGTCGCTT 274
QY 190 GTTCCGCGCAGTACGCAAGCTGCTTCAACAGCGGATCTACTTATATGAGCGCTCCTAGT 249
DB 275 GTTCCGCGCAGTACGCAAGCTGCTTCAACAGCGGATCTACTTATATGAGCGCTCCTAGT 334
QY 250 TGTAGTGGGATGGATGCTGCTACTTCCATGCAACGCTGAGTTTCCTGGGTCCAGATGCT 309
DB 335 TGTAGTGGGATGGATGCTGCTACTTCCATGCAACGCTGAGTTTCCTGGGTCCAGATGCT 394
QY 310 TGATGAATTCGCCATTCCTGCGGTTCTGATGTTGCTTTTGCCCAATGTTTCCAGAG 369
DB 395 TGATGAATTCGCCATTCCTGCGGTTCTGATGTTGCTTTTGCCCAATGTTTCCAGAG 454
QY 370 GTATTTACCAAGATCTTTTCGAATGACAGGCGAGTTTCAAGCGAGTGGTGTGTCCT 429
DB 455 GTATTTACCAAGATCTTTTCGAATGACAGGCGAGTTTCAAGCGAGTGGTGTGTCCT 514
QY 430 GTCTGCAATTACACGCTGCTTGGCGTTTATCAAGCCGCCATCAACAATATTTCCCTGAT 489
DB 515 GTCTGCAATTACACGCTGCTTGGCGTTTATCAAGCCGCCATCAACAATATTTCCCTGAT 574
QY 490 GATTCTGGGACTTCCATGCACTGCGCTGCTGTTGTCGAGAGCTGAAGAGTGTGCAATGT 549
DB 575 GATTCTGGGACTTCCATGCACTGCGCTGCTGTTGTCGAGAGCTGAAGAGTGTGCAATGT 634
QY 550 GCGTGTGTTTAAAGCTGGGCGCTTCTCTGGCGCTCTGGTGGACTCTT-GGCTCTCTTCTGCT 608
DB 635 GCGTGTGTTTAAAGCTGGGCGCTTCTCTGGCGCTCTGGTGGACTCTTGGGCTCTCTTCTGCT 694
QY 609 GGATCAGCGACCAAGCCCTTCTGTGAGCTGCTCTCTCTCTTTCACCTTCCCTACCTGCACT 668
DB 695 GGATCAGCGACCAAGCCCTTCTGTGAGCTGCTCTCTCTCTTTCACCTTCCCTACCTGCACT 754
QY 669 GTGTGTGGCATATTCTCATCTGCGCTTGTTCGTAACCTGGGCTGTGTGTGCTTCCGCTACT 728
DB 755 GTGTGTGGCATATTCTCATCTGCGCTTGTTCGTAACCTGGGCTGTGTGTGCTTCCGCTACT 814
QY 729 TTGATGTCGCTCAGAGATACCTGAGCAAGTCCAGT-CATCAGATTCTT-GGCGCCAGCGA 786

Db 815 TTGATGTCGCTCAGAGATACCTGAGCCAAAGTCCAGTCCATCAGATTCTGGGCCCCACCGA 874
QY 787 GAAATGGGC--TTTATTGGTGTCCTTATGTGTGTCCTTCTGTGTGCC-ACAAGAAGTC 843
Db 875 GGAATGGGCCTTTTATTGGGGGTCCTTATGGTGGTCCCTTCTGTGTGCCAACAGAAAGTC 934
QY 844 GCCAGTCAAGAT-CACGTGATGGCAAGGAGTGA--CCAGTCTTCTACTTACTTCTTATT 900
Db 935 GCCAGTCAAGATCCACGTGATGGCAAGGAGTGAACCAAGCGCTCCCTACTTACTTCCAAT 994

RESULT 4

CN537424

LOCUS

DEFINITION CN537424 717 bp mRNA linear EST 29-APR-2004
UI-M-HSO-cqo-o-10-0-UI.r1 NIH_BMAP_HSO Mus musculus cDNA clone
IMAGE:30672249 5', mRNA sequence.

ACCESSION

CN537424

VERSION

CN537424.1

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Chordata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 (bases 1 to 717)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: csapbs-2@mail.nih.gov

Tissue Procurement: Dr. James Lin University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uiowa.edu/distribution/mousefl.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..717

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:30672249"

/tissue_type="Upper Head"

/dev_stage="embryo 9.5 - 10.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_HSO"

/notes="Organ: Upper Head; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is CGACATGAAT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH)."

ORIGIN

Query Match 17.1%; Score 714.4; DB 7; Length 717;
Best Local Similarity 99.7%; Pred. No. 1e-185;
Matches 715; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 717 GCTTCGCTACTTTGATGCTGCCTCAGAGATACCTGAGCAAGGTCAGTTCATCAGATTCT 776
DB 1 GCTTCGCTACTTTGATGCTGCCTCAGAGATACCTGAGCAAGGTCAGTTCATCAGATTCT 60

Db 63 CAGTGTGTTGAGGATCAACAATGCGCTCTGTGAAATAATCTGCCACCCATTTCTATTGGT 122
 QY 3568 TGGATTCTCAGGAGTCTGAATCTTCCTCAGAGTCTCTCTCTCCCAACCCCTTACCAG 3627
 Db 123 TGGATTCTCAGGAGTCTGAATCTTCCTCAGAGTCTCTCTCTCCCAACCCCTTACCAG 182
 QY 3628 AGCAACACTGGGATTTGAACCTTCTCTGACTCTCTCTCTCCCTCAGAGTCTGACAACTAA 3687
 Db 183 AGCAACACTGGGATTTGAACCTTCTCTGACTCTCTCTCTCTCCCTCAGAGTCTGACAACTAA 242
 QY 3688 TGTGCTCTGGGGACACCCAGCTAGGGCTTCCCAACTCTCTTATCCAGCTGAACCTTCGAT 3747
 Db 243 TGTGCTCTGGGGACACCCAGCTAGGGCTTCCCAACTCTCTTATCCAGCTGAACCTTCGAT 302
 QY 3748 TCTTCCCAACCCAGGCTTGAACCTTCTCTGACTCTCTCTCTCTCCCTCAGAGTCTGACAACTAA 3807
 Db 303 TCTTCCCAACCCAGGCTTGAACCTTCTCTGACTCTCTCTCTCTCTCCCTCAGAGTCTGACAACTAA 362
 QY 3808 GGCAGATTGGGATTTCAATAGATTTTATAGACATTTATCTCCCAAGAGCTTTAAACA 3867
 Db 363 GGCAGATTGGGATTTCAATAGATTTTATAGACATTTATCTCCCAAGAGCTTTAAACA 422
 QY 3868 TGGCTTGTGCTTTCCATACACATCCGGTCTGATTTAAACTATTTTATTAACACAGAA 3927
 Db 423 TGGCTTGTGCTTTCCATACACATCCGGTCTGATTTAAACTATTTTATTAACACAGAA 482
 QY 3928 TTAACCAACCAATAGAGTCTTTCAGATATAAACTGTGTTTCATATTTATAGTAGAGT 3987
 Db 483 TTAACCAACCAATAGAGTCTTTCAGATATAAACTGTGTTTCATATTTATAGTAGAGT 542
 QY 3988 GTGCTATGATAGCGGTATGATCCCTGGCTGAAGTAAATTAACCATAGTCTGGGAGG 4047
 Db 543 GTGCTATGATAGCGGTATGATCCCTGGCTGAAGTAAATTAACCATAGTCTGGGAGG 602
 QY 4048 ATTACAGACCTTTTGACCTTTATGCTTTTGTGAACTCTGATAAACCAGTCAATATT 4107
 Db 603 ATTACAGACCTTTTGACCTTTATGCTTTTGTGAACTCTGATAAACCAGTCAATATT 662
 QY 4108 AAAGCCCAATAGTGGCAATTTCTGTG 4133
 Db 663 AAAGCCCAATAGTGGCAATTTCTGTG 688

RESULT 6
 AY401891
 LOCUS
 DEFINITION
 Mus musculus HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 AY401891
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 681)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 JOURNAL
 PUBLISHED
 14671302
 REFERENCE
 2 (bases 1 to 681)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 JOURNAL
 COMMENT
 This sequence was made by sequencing genomic exons and ordering

them based on alignment.
 Location/Qualifiers
 1..681
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..5681
 /locus_tag="HCM1041"
 ORIGIN
 Query Match 16.3% Score 681; DB 9; Length 681;
 Best Local Similarity 100.0%; Pred. No. 1.8e-176; Indels 0; Gaps 0;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 182 ATGTGCTTGTTCGCCAGTACGCAACGCTGCTTCAACAGCGGCATCTACTTAAATATGAGC 241
 Db 1 ATGTGCTTGTTCGCCAGTACGCAACGCTGCTTCAACAGCGGCATCTACTTAAATATGAGC 60
 QY 242 CTCCTAGTGTAGTGGGATTTGGATCTGTCTACTTCCATGCAAGCTGAGTTCTCTGGT 301
 Db 61 CTCCTAGTGTAGTGGGATTTGGATCTGTCTACTTCCATGCAAGCTGAGTTCTCTGGT 120
 QY 302 CAGATGCTTCATGAACCTTGCATTTCTGTGGTCTCTGATGTGTCTTTGGCCATGTGGTT 361
 Db 121 CAGATGCTTCATGAACCTTGCATTTCTGTGGTCTCTGATGTGTCTTTGGCCATGTGGTT 180
 QY 362 CCCAGGAGGTATTTACCAAGATCTTTTCGGAATGACAGGGGCGAGTTTCAAGGCGAGTGGT 421
 Db 181 CCCAGGAGGTATTTACCAAGATCTTTTCGGAATGACAGGGGCGAGTTTCAAGGCGAGTGGT 240
 QY 422 TGTGCTCTGTCTGCAATTAACAGTGTCTTGGCGTTTATCAAGCCCGGCATCAACAATATT 481
 Db 241 TGTGCTCTGTCTGCAATTAACAGTGTCTTGGCGTTTATCAAGCCCGGCATCAACAATATT 300
 QY 482 TCCCTGATGATTTCTGGGACTTCCATGCACTGCGTCTGTTGAGAGCTGGAAGAGTGT 541
 Db 301 TCCCTGATGATTTCTGGGACTTCCATGCACTGCGTCTGTTGAGAGCTGGAAGAGTGT 360
 QY 542 GACAAATGTGCTGTGTTTAAAGCTGGGCTCTTCTCTGGCCCTCTGTGGAGCTCTGGCTCTC 601
 Db 361 GACAAATGTGCTGTGTTTAAAGCTGGGCTCTTCTCTGGCCCTCTGTGGAGCTCTGGCTCTC 420
 QY 602 TTCTGCTGGATCAGCGACCAAGCCCTCTGTGAGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 661
 Db 421 TTCTGCTGGATCAGCGACCAAGCCCTCTGTGAGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 480
 QY 662 CTGCACTGTGTGGCATAATTTCTCATCTGCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 721
 Db 481 CTGCACTGTGTGGCATAATTTCTCATCTGCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTT 540
 QY 722 GCCTACTTTGATGCTGCTCAGAGATACCTGAGCAAGTCCAGTCATCAGATTCTGGGCC 781
 Db 541 GCCTACTTTGATGCTGCTCAGAGATACCTGAGCAAGTCCAGTCATCAGATTCTGGGCC 600
 QY 782 AGCGAGAAATGGGCTTTTATTTGGTGTCCCTTATGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 841
 Db 601 AGCGAGAAATGGGCTTTTATTTGGTGTCCCTTATGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 660
 QY 842 TCGCAGTCAAGATCAGCTGA 862
 Db 561 TCGCAGTCAAGATCAGCTGA 681
 RESULT 7
 CF110927
 LOCUS
 DEFINITION
 Shultzomica04178 Rat lung airway and parenchyma cDNA libraries
 Rattus norvegicus cDNA clone Contigs708 5', mRNA sequence.
 CF110927
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 CF110927.1 GI:33167279
 EST.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 1388)

REFERENCE
AUTHORS Shultz, M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V.,
Padua, A.M., Gurske, W.A., Morin, D., Penn, S.G., Jovanovich, S.B.,
Plopper, C.G. and Buckpitt, A.R.
TITLE Gene expression analysis in response to lung toxicants: I.
JOURNAL Sequencing and microarray development
COMMENT Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698
Email: mashultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 1388.

FEATURES
source Location/Qualifiers
1..1388
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="Contig3708"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cdna libraries"
/notes="Organ: lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."

ORIGIN
Query Match 16.1%; Score 670.6; DB 6; Length 1388;
Best Local Similarity 87.2%; Pred. No. 1.7e-173;
Matches 854; Conservative 0; Mismatches 104; Indels 21; Gaps 10;
QY 3141 CCACAGCGAAGCCCGTAAACCGATGCACTCCAGAACTCTTCCCTCTGCTGAGAGAT 3200
DB 419 CAAGAAGCAGAGCAGCGTAACTGATGCACTCCAGAACTCTGCTCCCTCCGCTGAGAT 478
QY 3201 GACATCGATGATGATGCTGTCACCACTGGCCCTGCATGGGTCTGCTCTCTCTGT 3260
DB 479 GACCTCAGTGATGATGCTGATGGCCACTGTTCCAGCATGGCGTCAGGTCTGCTCTGT 538
QY 3261 TCCGAATCTACCTGAGATCTCAGGACAGAGGAGCCATGAATGTTACCAAGTGGTCATGG 3320
DB 539 TCCGAATCTGCTGGATCTCAGGACAGAGGAGCCATGAAGTTACCAAGTGGTCATGT 598
QY 3321 CTGTCAGTGAATTTAGATTTTGAACATATTTGGTTTATAGA---GAATCTCTCTCTC 3377
DB 599 CTGTCAGTGAATTTAGTCTTTGATCCATTTGATTTTTCAGAGAGGAATTTCTCTCTC 658
QY 3378 TCTAGTGCCTGTGATGCAAGCCAGCCCTCAGAAAGTCTCTCTCTCTCTCTCTCTCT 3437
DB 659 TTTAGTGCCTGTGATGCAAGCCAGCCCTCAGAAAGTGTCTTCTGCTCTCTCTCTCT 718
QY 3438 TTGTAATGTGATCCAC - TCGGGGAAATGGGTGTCTACCCAGGAGAAAGTCCCTACCCAG 3496
DB 719 TTGTAATGTGATCCAGTGGGGAAATGGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 778
QY 3497 TCACGGGTCAACAGTGTGTGTGTGAGGATCAACATGCTCTGTGAAATATCTGCCACCC 3556
DB 779 TCACGGGTCAACAGCATTTGTGTGAGGATCAACATGCTCTGTGAAATATCTGCCACCC 838
QY 3557 ATTCTATTGGTGTGATTTCTCAGGAGTCTGAATCTTCCCTCAGAGTCTCTCTCTCTCT 3616
DB 839 ATTCTATTGGTGTGATTTCTCAGGAGTCTGAATCTTCCCTCAGAGTCTCTCTCTCT 898
QY 3617 CCCCTACCCAGAGCCCAACTGGGATTTGAACCTTC---TCTGACTCTCTCTCTCTCTCT 3672

DB 899 CGCCCTCCCGATGTC---CACTGGGATTGAACCTTCAAAATCCCTCTCTCTCTCTCT 955
QY 3673 AGGTCTGACAACTAATGGTCTCTGGGGACACCCAGCTAGGGCTTCCCCAACTCTCT-AT 3731
DB 956 AGGTCTGACCACTAATGGTCTCTGGGGACACCTAATAGGGCTTCCCCAACTCTCTCTCT 1015
QY 3732 CGAGCTGAATGGATTTCTTCCCAACAGGGCTTGAACCTGGGTCTGGTCCCACTGG 3791
DB 1016 CCAGCCAAATTTGGATTTCTTCCCAACAGGGCTTGAACCTGGGTCTGGTCCCACTGC 1075
QY 3792 CCAACAACACATCTTTGGCCAGATTTGGGATTTCTCAATAGATTTTATAGACATTTCTCC 3851
DB 1076 CCAATGACACATCTTTGACACAGATTTGGGATTTCTCAATAGATTTTATAGACATTTCTCC 1135
QY 3852 CACAGACTTT--AAAACATGGCTTGTCTTTTCCATACACATCCGGTTCAGATTTAAACT 3909
DB 1136 CACAGACTTTTAAACATGGCTTGTCTTTTCTTACACA---AGTCAGATTTAAACT 1192
QY 3910 ATTTTATACCAAGGAATTAACCAAGCAATAGAGTACTTTCAGATATAAACTGTGT 3969
DB 1193 ATTTTATACCAAGGAATTAACCAAGCAATAGAGTACTTTCAGATATAAACTGTGT 1250
QY 3970 TCATCTTTATGATGATGCTGTATGTATAGGGGGTATGTACCTGCTGCTGAAGTAAATATT 4029
DB 1251 TTATCTTTATGATGATGCTGTATGTATAGGGGGTATGTACCTGCTGCTGAAGTAAATATT 1310
QY 4030 AACCATAGCTCTGGGAGGATTTACAGACCTTTTTCACCTTTATGCT-TTTTTGTGAATCT 4088
DB 1311 AGTCATTGCTCTGGG-GGATTTACAGATCTTTTGCATTTATGCTTTTGTGAACTCT 1369
QY 4089 GATACCATGCTCAATATT 4107
DB 1370 GATAACCATGTTCAATATT 1388

RESULT 8
CN721804
LOCUS CN721804
DEFINITION E0830A03-5 NIA Mouse four-cell-Embryo cdna library (Long) Mus
musculus cdna clone NIA:E0830A03 IMAGE:30908450 5', mRNA sequence.
CN721804
ACCESSION CN721804.1 GI:47491189
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 670)
AUTHORS Sharov, A.A., Piao, Y., Matoba, R., Dudekula, D.B., Qian, Y.,
VanBuren, V., Falco, G., Martin, P., Stagg, C.A., Bassey, U.C.,
Wang, Y., Carter, M.G., Hamatani, T., Alba, K., Akutsu, H., Sharova, L.,
Tanaka, T.S., Kimber, W.L., Yoshikawa, T., Jaradat, S.A., Pantano, S.,
Nagaraja, R., Boheler, K.R., Taub, D., Hodes, R.J., Longo, D.L.,
Schlessinger, D., Kellier, J., Klotz, B., Kelsoe, G., Umezawa, A.,
Vescovi, A.L., Rossant, J., Kunath, T., Hogan, B.L., Curci, A.,
D'Urso, M., Kelsie, J., Hide, W. and Ko, M.S.
Transcriptome analysis of mouse stem cells and early embryos
PLOS Biol. 1 (3), 410-419 (2003)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: E0830 row: A column: 03
Seq primer: M13 Reverse
High quality sequence stop: 670
POLYA=No.
Location/Qualifiers
1..670
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

FEATURES
source
1..670

```

/db_xref="niaEST:E0830A03-5"
/db_xref="taxon:10090"
/clone="NIA:E0830A03 IMAGE:30908450"
/tissue type="4-cell stage embryo"
/dev stage="4-cell"
/lab_host="DH10B"
/clone_lib="NIA Mouse four-cell-Embryo cDNA library
(long)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library [Ref].
Genome Res. 11: 1553-1558 (2001). [PMID: 1154199]. The
mRNAs were extracted from a pool of 360 embryos at 4-cell
stage. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen]:
5'-pGAGTGTCTAGATCGGAGCGGCCCTTTTCTTTT-3' from
10.8ng of mRNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lone-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.2Kb. The library was
constructed by Yulan Piao."

```

ORIGIN

```

Query Match      16.0%; Score 668.4; DB 7; Length 670;
Best Local Similarity 99.9%; Pred. No. 5.5e-173;
Matches 669; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2210 ATGACTTCATCTGCTTTTGGCTTTGAGTTTAAAGTAACTGCTGGGCTATTTC 2269
DB 1 ATGACTTCATCTGCTTTTGGCTTTGAGTTTAAAGTAACTGCTGGGCTATTTC 60

QY 2270 TTGTTTATAGTCACAAATATTATAGAACATGAGGTGAATAAGTGTCTTTTATTA 2329
DB 61 TTGTTTATAGTCACAAATATTATAGAACATGAGGTGAATAAGTGTCTTTTATTA 120

QY 2330 AATTATAGCATTTACCAACTCTCCAGGTAGCAACACACAGTCATTTGATTCGCGATT 2389
DB 121 AATTATAGCATTTACCAACTCTCCAGGTAGCAACACACAGTCATTTGATTCGCGATT 180

QY 2390 TACAGGAGCAATACCTGTTTGAATGATGTATATGTCATTTTACTGTAGCTTAAAA 2449
DB 181 TACAGGAGCAATACCTGTTTGAATGATGTATATGTCATTTTACTGTAGCTTAAAA 240

QY 2450 TATGTTTAAATGACTTTGAGTGAATGTTTGTGCAACCTAGGAGTTTATGATCAGAA 2509
DB 241 TATGTTTAAATGACTTTGAGTGAATGTTTGTGCAACCTAGGAGTTTATGATCAGAA 300

QY 2510 TTGCGGCTGAGTCCTTTGGTTTTTGGTCTAGAAAGTTTTCAGAGGCGAAATCAAACTAG 2569
DB 301 TTGCGGCTGAGTCCTTTGGTTTTTGGTCTAGAAAGTTTTCAGAGGCGAAATCAAACTAG 360

QY 2570 CTTTGTGTTGTTTCATCTTACCAAGCTCGGAGCTGCCCTTATATTTACATAAGGA 2629
DB 361 CTTTGTGTTGTTTCATCTTACCAAGCTCGGAGCTGCCCTTATATTTACATAAGGA 420

QY 2630 CTTTATTATACATAGTCTGTATTAATGTCCTGAAGATGACACCTAGCTGCTTATCTGG 2689
DB 421 CTTTATTATACATAGTCTGTATTAATGTCCTGAAGATGACACCTAGCTGCTTATCTGG 480

QY 2690 AAGGGTCTCTGGGGCTGGAGAGTTGGTTCAACAGTTTGAAGAACCGGTGTTCTCTCCC 2749
DB 481 AAGGGTCTCTGGGGCTGGAGAGTTGGTTCAACAGTTTGAAGAACCGGTGTTCTCTCCC 540

```

```

QY 2750 AGACGACCCAGCGACTTGTGAGGCTCATATACAGCAACTCCAGCCCTGGGGCATCTGACG 2809
DB 541 AGACGACCCAGCGACTTGTGAGGCTCATATACAGCAACTCCAGCCCTGGGGCATCTGACG 600

QY 2810 CCCTCTTCTGACTTCTTAAGGGCATTTGGTGCACACGTATACATACAGGCGAGCAAAACATTT 2869
DB 601 CCCTCTTCTGACTTCTTAAGGGCATTTGGTGCACACGTATACATACAGGCGAGCAAAACATTT 660

QY 2870 ATACATGTAA 2879
DB 661 ATACATGTAA 670

AK085306 1173 bp mRNA linear HTC 03-APR-2004
Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
library, clone:D63008P07 product:similar to CANCER RELATED
GENE-LIVER 1 [Mus musculus], full insert sequence.
AK085306
AK085306.1 GI:26351558
HTC; CAP trapper.
Mus musculus (house mouse)
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
93279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Hashiwa, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
REFERENCE
AUTHORS
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
REFERENCE
AUTHORS
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
MEDLINE
11076861
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
REFERENCE
AUTHORS
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
MEDLINE
11076861
6 (bases 1 to 1173)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, I., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

```

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/

FEATURES

source

Location/Qualifiers

1..1173

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:D630008P07"

/db_xref="taxon:10090"

/clone="D630008P07"

/tissue_type="kidney"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

72..731

/note="unnamed protein product; putative

similar to CANCER RELATED GENE-LIVER 1 [Mus musculus]

(SPTK1AAL40408, evidence: FASTY, 100%ID, 77.8%length,

match=642)"

/codon_start=1

/protein_id="BAC39416.1"

/db_xref="GI:26315159"

/translation="MQAHFWHDLRAGSEVDWEDNYTIVPAIFNTISNVLPFI
LPIICMCFPRATCSGIYLIWLLVVGISGVVFHATLSFLQMDLDEAILVLM
CALAWMPFRYLPKIFRNDGRFKAIVCVVSAITTCFLAFKPAINISIMILGLPCTA
LLVABELKCDNRVRFKGLGSLWLTALFCWISDQAFCELLSSFFPYLHCWVSADR
G"

ORIGIN

Query Match 16.0%; Score 667.6; DB 3; Length 1173;
Best Local Similarity 98.7%; Pred. No. 1.1e-172;
Matches 673; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 10 GCTGCTCCGATGCTCCAGAGCGGCATGGCGCCGCGACTGGTGGGACCACTGGCGG 69
DB 47 GCTGCTCCGATGCTCCAGAGCGGCATGGCGCCGCGACTGGTGGGACCACTGGCGG 106
QY 70 TGGCAGTTCCGAGGTGGATTGGTGGGAGGACCACTACACTATCGTGGCTGCCA 129
DB 107 TGGCAGTTCCGAGGTGGATTGGTGGGAGGACCACTACACTATCGTGGCTGCCA 166
QY 130 GTTCTACACAGATCAGCAAGCTCTGTTTTCATTTTACCTCCCATCTGATGCTT 189
DB 167 GTTCTACACAGATCAGCAAGCTCTGTTTTCATTTTACCTCCCATCTGATGCTT 226
QY 190 GTTCCGCCAGTACGCAACGTCTTCAACAGCGGCATCTACTTAATATGACGCTCTACT 249
DB 227 GTTCCGCCAGTACGCAACGTCTTCAACAGCGGCATCTACTTAATATGACGCTCTACT 286
QY 250 TGTAGTGGGATGGATCTGTCTACTTCCATCCACAGCTGAGTTCTCTGGGTGAGATGCT 309
DB 287 TGTAGTGGGATGGATCTGTCTACTTCCATCCACAGCTGAGTTCTCTGGGTGAGATGCT 346
QY 310 TGATGACTTGCATCTCTGGGTCTGATGTCGTTGGCCATGCTGTTCCAGGAG 369
DB 347 TGATGACTTGCATCTCTGGGTCTGATGTCGTTGGCCATGCTGTTCCAGGAG 406

QY 370 GTATTTACCAAGATCTTTCCGAATGACAGGGGACAGGTTCAAGCAGTGTGTGTGTCCT 429
DB 407 GTATTTACCAAGATCTTTCCGAATGACAGGGGACAGGTTCAAGCAGTGTGTGTGTCCT 466
QY 430 GTCCTCAATTTACCAAGCTGCTTGGCGTTTATCAAGCCCGCCATCAACAATATTTCCCTGAT 489
DB 467 GTCCTCAATTTACCAAGCTGCTTGGCGTTTATCAAGCCCGCCATCAACAATATTTCCCTGAT 526
QY 490 GATTTCTGGGACATTTCCATGCACTGCGCTGCTTGTGTGAGAGCTCAAGAGGTGTGACAAATGT 549
DB 527 GATTTCTGGGACATTTCCATGCACTGCGCTGCTTGTGTGAGAGCTCAAGAGGTGTGACAAATGT 586
QY 550 GCGTGTGTTTAACTGCGGCTCTTCTGCGCTCTGCTGAGACTCTGCTGCTCTTCTGCTG 609
DB 587 GCGTGTGTTTAACTGCGGCTCTTCTGCGCTCTGCTGAGACTCTGCTGCTCTTCTGCTG 646
QY 610 GATCAGGACCAAGCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 669
DB 647 GATCAGGACCAAGCTTCTGTGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 706
QY 670 TGTGTGGCATATTTCTCATCTGC 691
DB 707 TGTGTGGAGTGTGCTGACAGGCG 728

RESULT 10

AV227941

LOCUS

DEFINITION

musculus cDNA clone 443241104 3', mRNA sequence.

ACCESSION

AV227941

VERSION

EST.

KEYWORDS

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 955)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,

Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,

Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,

Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,

Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

On Nov 1, 1999 this sequence version replaced gi:6180460.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.riken.jp. URL: http://genome.gsc.riken.jp/

Carninci, P., Shibata, Y., Hayashizaki, Y., Suganuma, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Suganuma, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a


```
QY 24 CCAGAGCGGCATGGGCCCCCGACCTGGTGGGACCACTGCGGCGTGGAGTTTCGGAGG 83
Db 1 CCAGAGCGGCATGGGCCCCCGACCTGGTGGGACCACTGCGGCGTGGAGTTTCGGAGG 60
QY 84 TGGATTGGTGGGAGGACAACTACACTATGCTGCTGCATGCGGAGTTCTCAACAACGA 143
Db 61 TGGATTGGTGGGAGGACAACTACACTATGCTGCTGCATGCGGAGTTCTCAACAACGA 120
QY 144 TCAGCAACGCTCTGTTTTCATTTTACCTCCCATCTGCATGCTGCTTCCGCGCAGTACG 203
Db 121 TCAGCAACGCTCTGTTTTCATTTTACCTCCCATCTGCATGCTGCTTCCGCGCAGTACG 180
QY 204 CAACGCTCTTCAACAGCGGCATCTACTTAATATGACGCTCCTAGTTGTAGTGGGATTG 263
Db 181 CAACGCTCTTCAACAGCGGCATCTACTTAATATGACGCTCCTAGTTGTAGTGGGATTG 240
QY 264 GATCTGCTACTTCCATGCAAGCGTGAATTCCTTGGTCAAGTCTTGATGAATTCGCCA 323
Db 241 GATCTGCTACTTCCATGCAAGCGTGAATTCCTTGGTCAAGTCTTGATGAATTCGCCA 300
QY 324 TTCTGTGGGTTCTGATGTGCTTTGGCCATGTGGTTTCCAGAGGATTTTACCAAGA 383
Db 301 TTCTGTGGGTTCTGATGTGCTTTGGCCATGTGGTTTCCAGAGGATTTTACCAAGA 360
QY 384 TCCTTCGGAATGACAGCGGCAGTTCAAGCAGTGTGTGTCCTGTCTGCAATTACAA 443
Db 361 TCCTTCGGAATGACAGCGGCAGTTCAAGCAGTGTGTGTCCTGTCTGCAATTACAA 420
QY 444 CGTCTGTGGGTTTATCAAGCCGCCATCAACAATATTTCCCTGATTTCTGGGACTTC 503
Db 421 CGTCTGTGGGTTTATCAAGCCGCCATCAACAATATTTCCCTGATTTCTGGGACTTC 480
QY 504 CATGCACTGCGCTGCTTTGTCAGAGCTGAAGAGGTGTGACAATGTGCGTGTGTTAAGC 563
Db 481 CATGCACTGCGCTGCTTTGTCAGAGCTGAAGAGGTGTGACAATGTGCGTGTGTTAAGC 540
QY 564 TGGGCTCTTCTCTGGCTCTGTGTGACTGTGCTCTCTTCTGCTGATGATGAGCAAG 623
Db 541 TGGGCTCTTCTCTGGCTCTGTGTGACTGTGCTCTCTTCTGCTGATGATGAGCAAG 600
QY 624 CCTTCTGTGAGCTCTCTCTCTTTCACCTTCCCTACCTGCACTGTGTGCAATTC 683
Db 601 CCTTCTGTGAGCTCTCTCTCTTTCACCTTCCCTACCTGCACTGTGTGAGTGTGCTG 660
QY 684 TCATCTGC 691
Db 661 ACAGAGGC 668
```

```
RESULT 12
CA463294
LOCUS AGENCOURT_10691314 NIH_MGC_169 Mus musculus cDNA clone
DEFINITION IMAGE:6770649 5', mRNA sequence.
ACCESSION CA463294
VERSION CA463294.1 GI:24919646
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 797)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
cDNA Library Prepared by: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
Plate: LICM3088 row: p column: 08
High quality sequence stop: 536.
FEATURES
Location/Qualifiers
source 1..797
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6770649"
/lab_host="DHI0B (Tl-phage-resistant)"
/notes="Organ: Testicles; Vector: pDNR-LIB; Site 1: Sfil
(ggcatatggcc); Site 2: Sfil (ggcgcctggcc); cDNA made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AACAGTGGTATCAACGAGTGGCCATTACGCGGG-3' and
5'-ATTCTAGAGCGGCGGCACATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN
Query Match 15.1%; Score 631.4; DB 6; Length 797;
Best Local Similarity 99.5%; Pred. No. 1e-162;
Matches 654; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 10 GCTGCTCCGATGCTCAGAGCGGCATGGGCGCCCGACCTGTGGAGCAACCTGGGGC 69
Db 10 GCTGCTCCGATGCTCAGAGCGGCATGGGCGCCCGACCTGTGGAGCAACCTGGGGC 69
QY 70 TGGCAGTTGGAGTGGATTGGTGGAGGACAACTACACTATCTGCTGCCATTGCCGA 129
Db 70 TGGCAGTTGGAGTGGATTGGTGGAGGACAACTACACTATCTGCTGCCATTGCCGA 129
QY 130 GTTCTACAACAGCATCAGCAACGCTTGTGTTTTCATTTTACCTCCCATCTGCAATGTCCT 189
Db 130 GTTCTACAACAGCATCAGCAACGCTTGTGTTTTCATTTTACCTCCCATCTGCAATGTCCT 189
QY 190 GTTCCGCGCAGTACGCAACGCTTCAACAGCGGCATCTACTTAATATGAGCGCTCTAGT 249
Db 190 GTTCCGCGCAGTACGCAACGCTTCAACAGCGGCATCTACTTAATATGAGCGCTCTAGT 249
QY 250 TGTAGTGGGATTTGGATCTGTCTACTTCCATCAACGCTGAGTTTCTGCGTCCAGATGCT 309
Db 250 TGTAGTGGGATTTGGATCTGTCTACTTCCATCAACGCTGAGTTTCTGCGTCCAGATGCT 309
QY 310 TGATGAACCTTGCCATTCTGTGGGTTCTGATGTGTGCTTTGGCCATGTGGTTTCCAGGAG 369
Db 310 TGATGAACCTTGCCATTCTGTGGGTTCTGATGTGTGCTTTGGCCATGTGGTTTCCAGGAG 369
QY 370 GTATTACCAAGATCTTTCGGAATGACAGGGGCGAGTTCAAGGAGTGGTGTGTCCT 429
Db 370 GTATTACCAAGATCTTTCGGAATGACAGGGGCGAGTTCAAGGAGTGGTGTGTCCT 429
QY 430 GTCTGCAATTACAAAGCTGTGCGTTTATCAAGCCCGCCATCAACAATATTTCCCTGAT 489
Db 430 GTCTGCAATTACAAAGCTGTGCGTTTATCAAGCCCGCCATCAACAATATTTCCCTGAT 489
QY 490 GATTCTGGGACTTCCATGCACTGCGCTGTGTTGTCAGAGCTGAAGAGGTGTGACAATGT 549
Db 490 GATTCTGGGACTTCCATGCACTGCGCTGTGTTGTCAGAGCTGAAGAGGTGTGACAATGT 549
QY 550 GCGTGTGTTTAAAGCTGGGCTCTTCTGCGCTCTGGTGGAGCTGCGCTCTCTCTGCTG 609
Db 550 GCGTGTGTTTAAAGCTGGGCTCTTCTGCGCTCTGGTGGAGCTGCGCTCTCTCTGCTG 609
QY 610 GATCAGCGCAACGAG-CCTTCTGTGAGCTGCTCT-CCTCTTTTCACTTCCCTTACCTG 664
Db 610 GATCAGCGCAACGAGCCCTTCTGTGAGCTGCTCTCCCTCTCTTCACTTCCCTTACCTG 666
RESULT 13
```


REFERENCE 1 (bases 1 to 617)
 AUTHORS Ronaldo.M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mES@mail.nih.gov
 Tissue Procurement: Dr. Xin-Yuan Fu, Yale University School of Medicine
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.
 Location/Qualifiers
 1..617
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CG0p-Bfj-a-09-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP Ret4 S2"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH-BMAP Ret4 S2 library is a subtracted library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine"

FEATURES
 source
 1..617
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CG0p-Bfj-a-09-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP Ret4 S2"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH-BMAP Ret4 S2 library is a subtracted library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine"

ORIGIN
 Query Match 14.2%; Score 594; DB 5; Length 617;
 Best Local Similarity 99.5%; Pred. No. 2.1e-152;
 Matches 616; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
 1584 GTGACCTCTGTGAGACAGGCGCTCGTTGTAATCGCGGAGGCGAGCTCTGAGGCTTGACT 1643
 Db 1 GTGACCTCTGTGAGACAGGCGCTCGTTGTAATCGCGGAGGCGAGCTCTGAGGCTTGACT 60
 1644 GTGGAGCGCTGTGAGATTTGGAAAGCTTTTCTATTGGAGCTTTGGTGTGCTT 1703
 Db 61 GTGGAGCGCTGTGAGATTTGGAAAGCTTTTCTATTGGAAAGCTTTGGTGTGCTT 120
 1704 ACCTCAAGATCTGACCCCGCTTCTCTGCTGTTTTAGGGATATGTCTTCTCGAGTG 1763
 Db 121 ACCTCAAGATCTGACCCCGCTTCTCTGCTGTTTTAGGGATATGTCTTCTCGAGTG 180
 1764 ACCTAGCCTCCCTAGATGTGGCCCTCCGGCTTTGTTTTCATPAACCTGGGTAACT 1823
 Db 181 ACCTAGCCTCCCTAGATGTGGCCCTCCGGCTTTGTTTTCATPAACCTGGGTAACT 240
 1824 GGTTCCTCAGACACTTCTCTCTGATGCTGGGGGTCTACTACACTGATGGCTG 1883
 Db 241 GGTTCCTCAGACACTTCTCTCTGATGCTGGGGGTCTACTACACTGATGGCTG 300
 1884 GCTGGGGATCCCAGGAGGAAGGGGAGACTCTGACATGAATGCTCTCACTGCATCCT 1943
 Db 301 GCTGGGGAT-CCAGGAGGAAGGGGAGACTCTGACATGAATGCTCTCACTGCATCCT 359
 1944 ACTGCTTCTACTGGGCTCTCTTCAAGGGTATGAAGTGGGTATGGGGTATCTCAGGATGTT 2003
 Db 360 ACTGCTTCTACTGGGCTCTCTTCAAGGGTATGAAGTGGGTATGGGGTATCTCAGGATGTT 419

QY 2004 TGTAACTCAGGACCTTCTGCTTTCTGACATTCATTCAGTGTGAGACTGCACCTCAG 2063
 Db 420 TGTAACTCAGGACCTTCTGCTTTCTGACATTCATTCAGTGTGAGACTGCACCTCAG 479
 QY 2064 GGGACTGGAGTTTGGAAACACCGTTCTGAGTGTGCTGAGTCTGAGAGGAGCTAGTCGGG 2123
 Db 480 GGGACTGGAGTTTGGAAACACCGTTCTGAGTGTGCTGAGTCTGAGAGGAGCTAGTCGGG 539
 QY 2124 GTTCTGGCACTTCTAGGATCTCTTACTCTGTTTGTAGAACCTTACAGGGTACAAAGTGGGA 2183
 Db 540 GTTCTGGCA-TTCTAGGATCTCTTACTCTGTTTGTAGAACCTTACAGGGTACAAAGTGGGA 598
 QY 2184 ACTGGACTTAAGAGTTTT 2202
 Db 599 ACTGGACTTAAGAGTTTT 617

RESULT 15
 AZ411514/c
 LOCUS
 DEFINITION
 AZ411514
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 606)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
 Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
 Niederhausern,A. and Wright,D.,Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0184 row: B column: 07
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 606.
 Location/Qualifiers
 1..606
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0184B07"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative

FEATURES
 source
 1..606
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0184B07"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN		Query Match	13.8%	Score 576.8	DB 8	Length 606
		Best Local Similarity	99.2%	Pred. No. 1.2e-147		
		Matches 601	Conservative 0	Mismatches 2	Indels 3	Gaps 2
Qy	1066	GTATCCTTTCCAGGCATACGGGCTTGCTGGGTTATGTCATATGACATCAACAGAGTAG	1125			
Db	606	GTATCCTTTCCAGGCATACGGGCTTGCTGGGTTATGTCATATGACATCAACAGAGTAG	547			
Qy	1126	TTCTTTGGGAGACCTAGGGCAACCAAGTTTCTTGCTGGAGAGGCTTGCTGCTTTC	1185			
Db	546	TTCTTTGGGAGACCTAGGGCAACCAAGTTTCTTGCTGGAGAGGCTTGCTGCTTTC	487			
Qy	1186	CACCATTTCCAGACTCTCAACCCCTGATAAACAAGACCTTCTGATTTGGTGATGAAAGG	1245			
Db	486	CACCATTTCCAGACTCTCAACCCCTGATAAACAAGACCTTCTGATTTGGTGATGAAAGG	427			
Qy	1246	TTCCAGAACTTTTCATTTTGGCGGAGAACTGCTTCAACAAAACCAAGTGGGCAAAA	1305			
Db	426	TTCCAGAACTTTTCATTTTGGCGGAGAACTGCTTCAACAAAACCAAGTGGGCAAAA	367			
Qy	1306	CACGTGTGGGGGTGTGACTAAGACGGGTGGCTTGTCACTCGCACCTCTTAGCTTTCCCA	1365			
Db	366	CACGTGTGGGGGTGTGACTAAGACGGGTGGCTTGTCACTCGCACCTCTTAGCTTTCCCA	307			
Qy	1366	GGTTCTCCAGTGTGTGGATCCTTCCACTACCTCTGCTGAGAGATGGAGCCACGGCTT	1425			
Db	306	GGTTCTCCAGTGTGTGGATCCTTCCACTACCTCTGCTGAGAGATGGAGCCACGGCTT	247			
Qy	1426	CAGAGGCAAAAGCTGGCAACACCCCTCTATGCCAAAAGTACACTCTCTTTAAGCACACA	1485			
Db	246	CAGAGGCAAAAGCTGGCAACACCCCTCTATGCCAAAAGTACACTCTCTTTAAGCACACA	187			
Qy	1486	TTAC--ATAGACATTTTCTGCTCTTCCAGAGTGCAGAGCCCTCAGACCCACACAGAGAT	1543			
Db	186	TTACATATAGACATTTTCTGCTCTTCCAGAGTGCAGAGCCCTCAGACCCACACAGAGAT	127			
Qy	1544	CCTTCAGGTTATGTAAGATTCCACACACAGCCCCCTCTTGTGACCTCTGTGAGAGCAAG	1603			
Db	126	CCTTCAGGTTATGTAAGATTCCACACACAGCCCCCTCTTGTGACCTCTGTGAGAGCAAG	67			
Qy	1604	GCCTCGTTGTAATCGCGAGGACGCTCTGGAGGCTTGACTGTGAGGCGCTGTGAAGATT	1663			
Db	66	GCCTCGTTGTAATCGCGAGGACGCTCTGGAGGCTTGACTGTGAGGCG-CTGTGAAGATT	8			
Qy	1664	TTGAAA 1669				
Db	7	TGAAA 2				

Search completed: November 21, 2004, 06:57:29
Job time : 11526.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 21, 2004, 03:07:33 ; Search time 3473 Seconds
(without alignments)
3744.505 Million cell updates/sec

Title: US-10-017-410-2

Perfect score: 1512
Sequence: 1 MGAPHWHDLRAGSSEVDWC.....IGVYVSVLLCAHKKSPVKIT 275

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgm2_1/USPTO.spool.p/US10017410/runat.19112004.133620.18996/app.query.fasta_1.455
-DB-GenEmbl -OFMT=fastad -SUFFIX=rge -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10017410 @CGN 1 1 5600 @runat.19112004.133620.18996 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb ba:*
2: gb bta:*
3: gb in:*
4: gb om:*
5: gb ov:*
6: gb pat:*
7: gb ph:*
8: gb pl:*
9: gb pr:*
10: gb ro:*
11: gb sts:*
12: gb sy:*
13: gb un:*
14: gb vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1512	100.0	4174	10 AF282864	Mus muscu
2	1426	94.3	2852	6 CQ842051	Sequence
3	1426	94.3	2852	9 AK123581	Homo sapi
4	1373.5	90.8	822	6 AX206799	Sequence

C	5	1361.5	90.0	149462	9	AC017081
	6	1252	82.8	4078	10	BC059819
	7	800.5	52.9	1429	3	AK116177
	8	714.5	47.3	2318	3	AY071232
	9	706.5	46.7	1163	3	AF323976
	10	549.5	36.3	792	6	AX206797
	11	549.5	36.3	795	9	AF347024
	12	547	36.2	315	6	CQ744096
	13	545	36.0	315	6	CQ736905
	14	545	36.0	315	6	CQ736906
	15	517.5	34.2	822	10	AF347023
	16	455	30.1	537	6	CQ605547
	17	340.5	22.5	90370	10	EX005084
	18	340.5	22.5	207360	2	EX324222
	19	336.5	22.3	112713	9	AL391834
	20	336.5	22.3	219679	2	AC109451
	21	334.5	22.1	184163	2	AC120245
	22	334.5	22.1	249734	2	AC097362
	23	320	21.2	4758	6	CQ605546
	24	320	21.2	13715	2	AC017332
	25	320	21.2	78857	3	AC004364
	26	320	21.2	157875	3	AE003665
	27	320	21.2	164361	3	AC009252
	28	320	21.2	165585	3	AC093044
	29	316.5	20.9	183542	9	AL158206
	30	316.5	20.9	219679	2	AC109451
	31	264.5	17.5	4071	6	CQ615131
	32	264.5	17.5	38138	2	AC017406
	33	227.5	15.0	677	5	CR523187
	34	219	14.5	257	6	CQ731034
	35	212.5	14.1	1313	8	AK103953
	36	212.5	14.1	1335	8	AK073303
	37	209.5	13.9	801	6	AX206798
	38	208.5	13.8	1347	9	BC073853
	39	208.5	13.8	3404	9	AF214454
	40	208.5	13.8	3473	9	AF327353
	41	204	13.5	358	6	AR419744
	42	204	13.5	358	6	AX980438
	43	204	13.5	358	6	BD115297
	44	204	13.5	799	8	BT008549
	45	204	13.5	1193	8	AY090947

ALIGNMENTS

RESULT 1	AF282864	Mus musculus cancer related gene-liver 1 mRNA, complete cds.	4174 bp	linear	ROD 12-DEC-2001
LOCUS	AF282864				
DEFINITION	AF282864				
ACCESSION	AF282864.1	GI:17529683			
VERSION					
KEYWORDS					
SOURCE	Mus musculus (house mouse)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 4174)				
AUTHORS	Graveel,C.R., Jatkoa,T., Madore,S.J., Holt,A.L. and Farnham,P.J.				
TITLE	Expression profiling and identification of novel genes in hepatocellular carcinomas				
JOURNAL	Oncogene 20 (21), 2704-2712 (2001)				
MEDLINE	21313787				
PUBMED	11420682				
REFERENCE	2 (bases 1 to 4174)				
AUTHORS	Graveel,C.R., Jatkoa,T., Madore,S.J., Holt,A.L. and Farnham,P.J.				
TITLE	Identification of genes deregulated in murine hepatocellular carcinomas using oligonucleotide microarrays and representational difference analysis				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 4174)				
AUTHORS	Farnham,P.J. and Graveel,C.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-JUN-2000) Oncology, University of Wisconsin, 1400				

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525


```

ACCESSION AX206799
VERSION AX206799.1 GI:15394643
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hofmann,K. and Conradt,M.
TITLE Ceramidase
JOURNAL Patent: WO 0155408-A 6 02-AUG-2001;
Memorec Medical Research Cologne Stoffel GmbH (DE)
FEATURES
source
1..822
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1..63e-121 Length: 822
Score: 1373.50 Matches: 246
Percent Similarity: 94.91% Conserved: 15
Best Local Similarity: 89.45% Mismatches: 13
Query Match: 90.84% Indels: 1
DB: 6 Gaps: 1
US-10-017-410-2 (1-275) x AX206799 (1-822)
QY 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
Db 1 ATGGGGCCCGCCACTGGTGGGACCGAGCTGCAGCTGGCAGCTGGAGGTGGAGCTGGCG 60
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 61 GAGGACAACCTACACCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 41 LeuPhePheIleLeuProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
Db 121 TTATTTTTCATTTCACGCCCATCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180
QY 61 AsnSerGlyIleTyrIleLeuThrLeuLeuValValValGlyIleGlySerValTyr 80
Db 181 AACAGCGGCATCTACTTAATCTGG--CTCTTGGTTGTAGCGGAATTGGATCGGCTAC 237
QY 81 PheHisAlaThrIleuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
Db 238 TTCCATGCAACCTTAGTTTCTCGGGTCAGATGCTTGATGAACCTTCAGTCTCTTGGGT 297
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db 298 CTGATGTGTGCTTCGGTTCATGTGGTCCCGAGAGGTACTACCAAGATCTTTCGGGAT 357
QY 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThrThrCysLeuAla 140
Db 358 GACCAGGTAGGTTCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 417
QY 141 PheIleLysProAlaIleAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
Db 418 TTGTCAAGCTCCCATCAACATCTCTCTGATGACCTGGAGGTCTCTCGGTGCA 477
QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
Db 478 CTGCTCATCAGAGCTAAGAGGTGTGACAAATCGGTGTGTAAAGCTGGGCTCTTC 537
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTyrTrpIleSerAspGlnAlaPheCysGlu 200
Db 538 TCGGGCTCTCTGGTGGACCTGGGCTGTCTGCTGGATCAGTCAAGCCAGCTTCTCGGAG 597
QY 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeuIleCysLeu 220
Db 598 CTGCTGTATCTTCACTTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
QY 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db 658 GCTGCTACTCTGGCTGTGTATGCTTTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 241 GlnGlyProValIleArgPheTrpProSerGluLysTrpAlaPheIleGlyValProTyr 260
Db 718 CRAAGCCCTGTCAATCAATTCGCCAGCAGAGAAATGGCCCTTCATTTGTTGTTGCTTAT 777
QY 261 ValSerLeuLeuCysAlaHisLysLysSerProValLysIleThr 275
Db 778 GTGTCCCTCTGTGTGTGCTGCAACAGAAATCATCATCAAGACCAG 822
RESULT 5
AC017081/c
LOCUS AC017081 149462 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-470J24 from 2, complete sequence.
ACCESSION AC017081
VERSION AC017081.8 GI:18093316
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 149462)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 149462)
AUTHORS Nguyen,C., Doeber,A. and Kozlowicz,A.
TITLE The sequence of Homo sapiens BAC clone RP11-470J24
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 149462)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Genome Sequencing Center, Washington
MO 63108, USA
REFERENCE 4 (bases 1 to 149462)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Jan 9, 2002 this sequence version replaced gi:14165368.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0470J24
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

```


SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-310K15, 2000 bp overlap; the clone sequenced to the right is RP11-156A1, 2000 bp overlap. Actual start of this clone is at base position 190775 of RP11-310K15; actual end is at base position 28935 of RP11-156A1.

The sequence between 81601 and 81703 is covered only by a PCR product of clone DNA. Unresolved tandem repeat regions exist between 81548 and 8183. 126196 and 127426.

FEATURES

source

1. 149462

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-470J24"

/clone_lib="RPCI-11"

141. 282

/rpt_family="MIR"

581. 621

/rpt_family="TGGG)n"

965. 1037

/rpt_family="GC_rich"

1127. 1561

/note="match to EST AA056210 (NID:gl548612) zf62g06.sl"

1771. 2127

/note="match to EST AW614820 (NID:g7320006) hg80c08.xl"

2055. 2195

/rpt_family="MER1_type"

2216. 2419

/rpt_family="MER1_type"

2604. 2844

/rpt_family="MIR"

2991. 3324

/note="match to EST R35058 (NID:g791959) yhs6d09.r1"

3207. 3758

/note="match to EST AW452004 (NID:g6992780)"

3486. 3527

/rpt_family="AT_rich"

3501. 3745

/note="match to EST AA054208 (NID:gl545131) zf51g03.r1"

3981. 4017

/rpt_family="T-rich"

4231. 4273

/rpt_family="AT_rich"

5518. 5847

/note="match to EST R80994 (NID:g857275) y194c08.r1"

5606. 5633

/rpt_family="AT_rich"

5722. 5758

/rpt_family="AT_rich"

5802. 5831

/rpt_family="(TTTAA)n"

5808. 6116

/rpt_family="Alu"

6153. 6244

/rpt_family="MER1_type"

6965. 7000

/rpt_family="AT_rich"

7020. 7094

/rpt_family="MER1_type"

7982. 8006

repeat_region
8000. 8436
/rpt_family="AT_rich"
repeat_region
8437. 8540
/rpt_family="L2"
repeat_region
8541. 8696
/rpt_family="L1"
repeat_region
8722. 9684
/rpt_family="L2"
repeat_region
9378. 9422
/rpt_family="L1"
repeat_region
9685. 9973
/rpt_family="(TTG)n"
repeat_region
9974. 10197
/rpt_family="Alu"
repeat_region
10280. 10305
/rpt_family="L1"
repeat_region
10320. 10341
/rpt_family="(TATG)n"
repeat_region
10333. 10679
/rpt_family="AT_rich"
repeat_region
10681. 10757
/rpt_family="L1"
repeat_region
10758. 11039
/rpt_family="Alu"
repeat_region
11040. 12011
/rpt_family="L1"
repeat_region
11999. 12591
/rpt_family="L1"
repeat_region
12592. 12706
/rpt_family="L1"
repeat_region
12707. 12952
/rpt_family="Alu"
repeat_region
12966. 13140
/rpt_family="L1"
repeat_region
13139. 16061
/rpt_family="L1"
repeat_region
16063. 17107
/rpt_family="L1"
repeat_region
17108. 17161
/rpt_family="MaLR"
repeat_region
17163. 17468
/rpt_family="MER1_type"
repeat_region
17471. 17675
/rpt_family="MaLR"
repeat_region
17677. 17835
/rpt_family="L1"
repeat_region
17909. 18144
/rpt_family="L1"
repeat_region
18177. 18501
/rpt_family="L1"
repeat_region
18519. 20393
/rpt_family="L1"
repeat_region
20490. 20906
/rpt_family="L1"
repeat_region
21258. 21365
/rpt_family="MIR"
repeat_region
21353. 21393
/rpt_family="(T)n"
repeat_region
21366. 21684
/rpt_family="Alu"
repeat_region
21543. 21566
/rpt_family="AT_rich"
repeat_region
21685. 21819
/rpt_family="MIR"
repeat_region
22160. 22640
/rpt_family="MER2_type"
repeat_region
22878. 23228

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

9.78e-118

1361.50

94.91%

Length:

Matches:

Conservative:

149462

246

15

Score: 549.50 Matches: 103
Percent Similarity: 59.38% Conservative: 49
Best Local Similarity: 40.23% Mismatches: 102
Query Match: 36.34% Indels: 2
DB: 6 Gaps: 1

US-10-017-410-2 (1-275) x AX206797 (1-792)

QY 9 HisLeuArgAlaGly-SerSerGluValAspTrpCysGluAspAsnTyrThrIleValPr 28
Db 9 CATCTTCGCTATCAGAGCTCCGAGTGGAGTGGTGTGAGAGCAACTCCAGTACTCGGA 68
QY 28 oAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProIle 48
Db 69 GCTGGTGGCGAGTCTTACACAGCTTCTCCATATCCCTTCTTATCTTCGGGCGACT 128
QY 48 eCysMetCysLeuPheArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTr 68
Db 129 GATGATGCTCTGATGCACCCGCTATGCCAGAGCGCTCCCGCTACATTTACGTTGTCIG 188
QY 68 pThrLeuValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 88
Db 189 GGTCTCTTCCATGATCATAGGCTGTCTCCATGTATTTCCACATGACGCTCAGCTTCCT 248
QY 88 uGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTr 108
Db 249 GGCCGAGCTGCTGGAGAGATCCCATCTCTGCTGCTGGGCGAGTGGCTATGACATATG 308
QY 108 pPheProArgTyrLeuPheValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 128
Db 309 GATGATGCTCTGATGCACCCGCTATGCCAGAGCGCTCCCGCTACATTTACGTTGTCIG 188
QY 68 pThrLeuValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 88
Db 189 GGTCTCTTCCATGATCATAGGCTGTCTCCATGTATTTCCACATGACGCTCAGCTTCCT 248
QY 88 uGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTr 108
Db 249 GGCCGAGCTGCTGGAGAGATCCCATCTCTGCTGCTGGGCGAGTGGCTATGACATATG 308
QY 108 pPheProArgTyrLeuPheValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 128
Db 309 GATGATGCTCTGATGCACCCGCTATGCCAGAGCGCTCCCGCTACATTTACGTTGTCIG 188
QY 128 lValCysValLeuValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 88
Db 369 GGTCTCTTCCATGATCATAGGCTGTCTCCATGTATTTCCACATGACGCTCAGCTTCCT 248
QY 88 uGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTr 108
Db 249 GGCCGAGCTGCTGGAGAGATCCCATCTCTGCTGCTGGGCGAGTGGCTATGACATATG 308
QY 108 pPheProArgTyrLeuPheValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 128
Db 309 GATGATGCTCTGATGCACCCGCTATGCCAGAGCGCTCCCGCTACATTTACGTTGTCIG 188
QY 128 lValCysValLeuValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 88
Db 369 GGTCTCTTCCATGATCATAGGCTGTCTCCATGTATTTCCACATGACGCTCAGCTTCCT 248
QY 148 nileSerLeuMetIleLeuGlyLeuProCysThrAlaLeuValAlaGluLeuLysAr 168
Db 429 CTACGCCCTCAACAGCATTCGCCATCTCTACATCGTGTGCCAGAGTACAGGA 488
QY 168 gCysAspAsnValArgValPheLeuGlyLeuPheSerGlyLeuTrpThrLeuAl 188
Db 489 GACCAGCAATAAGGAGCTTCGGCACCTGATGAGTGGTGTGAGAGCAACTCCAGTACTCGGA 68

Alignment Scores: 4.89e-43 Length: 795
Pred. No.: 549.50 Matches: 103
Score: 59.38% Conservative: 49
Percent Similarity: 40.23% Mismatches: 102
Best Local Similarity: 36.34% Indels: 2
Query Match: 9 Gaps: 1
DB: 1

US-10-017-410-2 (1-275) x AF347024 (1-795)

QY 9 HisLeuArgAlaGly-SerSerGluValAspTrpCysGluAspAsnTyrThrIleValPr 28
Db 9 CATCTTCGCTATCAGAGCTCCGAGTGGAGTGGTGTGAGAGCAACTCCAGTACTCGGA 68
QY 28 oAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProIle 48
Db 69 GCTGGTGGCGAGTCTTACACAGCTTCTCCATATCCCTTCTTATCTTCGGGCGACT 128
QY 48 eCysMetCysLeuPheArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTr 68
Db 129 GATGATGCTCTGATGCACCCGCTATGCCAGAGCGCTCCCGCTACATTTACGTTGTCIG 188
QY 68 pThrLeuValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 88
Db 189 GGTCTCTTCCATGATCATAGGCTGTCTCCATGTATTTCCACATGACGCTCAGCTTCCT 248
QY 88 uGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTr 108
Db 249 GGCCGAGCTGCTGGAGAGATCCCATCTCTGCTGCTGGGCGAGTGGCTATGACATATG 308
QY 108 pPheProArgTyrLeuPheValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 128
Db 309 GATGATGCTCTGATGCACCCGCTATGCCAGAGCGCTCCCGCTACATTTACGTTGTCIG 188
QY 128 lValCysValLeuValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 88
Db 369 GGTCTCTTCCATGATCATAGGCTGTCTCCATGTATTTCCACATGACGCTCAGCTTCCT 248
QY 148 nileSerLeuMetIleLeuGlyLeuProCysThrAlaLeuValAlaGluLeuLysAr 168
Db 429 CTACGCCCTCAACAGCATTCGCCATCTCTACATCGTGTGCCAGAGTACAGGA 488
QY 168 gCysAspAsnValArgValPheLeuGlyLeuPheSerGlyLeuTrpThrLeuAl 188
Db 489 GACCAGCAATAAGGAGCTTCGGCACCTGATGAGTGGTGTGAGAGCAACTCCAGTACTCGGA 68

RESULT 11
AF347024
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 795)
Mao, C., Xu, R. and Obeid, L.M.

```

Qy 188 aLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPhePr 208
|||||
Db 549 TCTGACCAGCTGGATCAGTACCGCTTCGTTTGAGCTTCTGGCAGAGGATTCAATTTCTT 608
|||||
Qy 208 oTyLeuHisCysValTrpHisIleLeuLeuAlaSerTyTrLeuGlyCysValCy 228
|||||
Db 609 CTATCTGCACGATCTGGCATGTGCTCATCAGCATCACCTTCCCTTATGGCATGGTAC 668
|||||
Qy 228 sPheAlaTyPheAspAlaSerGluIleProGluGlnGlyProValIleArgPheTr 248
|||||
Db 669 CATGGCCTTGCTGATGCCAACTATGATGATGCCAGGTGAAACCCCTCAAGTCCGCTACTG 728
|||||
Qy 248 pProSerGluTyTrpAlaPheIleGlyValProTyTrValSerLeu 263
|||||
Db 729 GCCTCGGGACAGTTGGCCC---GTGGGGCTGCCCTACGTGGAATC 771
|||||

RESULT 12
LOCUS CO744096 315 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 30030 from Patent WO02068579.
ACCESSION CO744096
VERSION CO744096.1 GI:42358801
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 30030 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source 1. .315
Location/Qualifiers
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.87e-43 Length: 315
Score: 547.00 Matches: 95
Percent Similarity: 99.04% Conservative: 8
Best Local Similarity: 91.35% Mismatches: 1
Query Match: 36.18% Indels: 0
DB: 6 Gaps: 0

US-10-017-410-2 (1-275) x CO744096 (1-315)
Qy 172 ValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCys 191
:::
Db 1 ATGGGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCTGTTGGACCTGGCCCTGTTCTGC 60
|||||
Qy 192 TrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPheProTyTrLeuHis 211
|||||
Db 61 TGGATCAGTGACCGAGCTTTCTGGAGCTGTGTCATCTTCAACTTCCCTTACCTGCAC 120
|||||
Qy 212 CysValTrpHisIleLeuLeuAlaSerTyTrLeuGlyCysValCysPheAlaTyTr 231
|||||
Db 272 ValLysIleThr 275
|||||

US-10-017-410-2 (1-275) x CO736905 (1-315)
Qy 172 ValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCys 191
:::
Db 1 ATGGGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCTGTTGGACCTGGCCCTGTTCTGC 60
|||||
Qy 192 TrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPheProTyTrLeuHis 211
|||||
Db 61 TGGATCAGTGACCGAGCTTTCTGGAGCTGTGTCATCTTCAACTTCCCTTACCTGCAC 120
|||||
Qy 212 CysValTrpHisIleLeuLeuAlaSerTyTrLeuGlyCysValCysPheAlaTyTr 231
|||||
Db 121 TGCATGTGGCAGCATCTCTCATCTGCTTGTCTGCTTACCTGGGCTGTGTATGCTTTCCTAC 180
|||||
Qy 232 PheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPheTrpProSerGlu 251
|||||
Db 181 TTTGATGCTGCTCAGAGATTCTCTGAGCAAGGCCCTGTCTATCAAAATTCTGGCCCGAGGAG 240
|||||
Qy 252 LysTrpAlaPheIleGlyValProTyTrValSerLeuLeuCysAlaHisLysLysSerPro 271
|||||
Db 241 AAATGGGCTTCAATGTGGTGTCCCTATGTGTCCCTCTGTGTCCTCTGTGTCACCAAGAAATCATCA 300
|||||
Qy 272 ValLysIleThr 275
|||||

US-10-017-410-2 (1-275) x CO736906 (1-315)
Qy 172 ValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCys 191
:::
Db 1 ATGGGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCTGTTGGACCTGGCCCTGTTCTGC 60
|||||
Qy 192 TrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPheProTyTrLeuHis 211
|||||
Db 61 TGGATCAGTGACCGAGCTTTCTGGAGCTGTGTCATCTTCAACTTCCCTTACCTGCAC 120
|||||
Qy 212 CysValTrpHisIleLeuLeuAlaSerTyTrLeuGlyCysValCysPheAlaTyTr 231
|||||
Db 121 TGCATGTGGCAGCATCTCTCATCTGCTTGTCTGCTTACCTGGGCTGTGTATGCTTTCCTAC 180
|||||
Qy 232 PheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPheTrpProSerGlu 251
|||||
Db 181 TTTGATGCTGCTCAGAGATTCTCTGAGCAAGGCCCTGTCTATCAAGTCTTGGCCCAATGAG 240
|||||
Qy 252 LysTrpAlaPheIleGlyValProTyTrValSerLeuLeuCysAlaHisLysLysSerPro 271
|||||
Db 241 AAATGGGCTTCAATGTGGTGTCCCTATGTGTCCCTCTGTGTCCTCTGTGTCACCAAGAAATCATCA 300
|||||
Qy 272 ValLysIleThr 275
|||||

```

```

Db 301 GTCAAGATCAGC 312

RESULT 13
LOCUS CO736905 315 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 22839 from Patent WO02068579.
ACCESSION CO736905
VERSION CO736905.1 GI:42333763
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 22839 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source 1. .315
Location/Qualifiers
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.45e-43 Length: 315
Score: 545.00 Matches: 95
Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 91.35% Mismatches: 2
Query Match: 36.04% Indels: 0
DB: 6 Gaps: 0

US-10-017-410-2 (1-275) x CO736905 (1-315)
Qy 172 ValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCys 191
:::
Db 1 ATGGGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCTGTTGGACCTGGCCCTGTTCTGC 60
|||||
Qy 192 TrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPheProTyTrLeuHis 211
|||||
Db 61 TGGATCAGTGACCGAGCTTTCTGGAGCTGTGTCATCTTCAACTTCCCTTACCTGCAC 120
|||||
Qy 212 CysValTrpHisIleLeuLeuAlaSerTyTrLeuGlyCysValCysPheAlaTyTr 231
|||||
Db 121 TGCATGTGGCAGCATCTCTCATCTGCTTGTCTGCTTACCTGGGCTGTGTATGCTTTCCTAC 180
|||||
Qy 232 PheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPheTrpProSerGlu 251
|||||
Db 181 TTTGATGCTGCTCAGAGATTCTCTGAGCAAGGCCCTGTCTATCAAAATTCTGGCCCGAGGAG 240
|||||
Qy 252 LysTrpAlaPheIleGlyValProTyTrValSerLeuLeuCysAlaHisLysLysSerPro 271
|||||
Db 241 AAATGGGCTTCAATGTGGTGTCCCTATGTGTCCCTCTGTGTCCTCTGTGTCACCAAGAAATCATCA 300
|||||
Qy 272 ValLysIleThr 275
|||||
Db 301 GTCAAGATCAGC 312

RESULT 14
LOCUS CO736906 315 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 22840 from Patent WO02068579.
ACCESSION CO736906
VERSION CO736906.1 GI:42333767
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 22839 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source 1. .315
Location/Qualifiers
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.45e-43 Length: 315
Score: 545.00 Matches: 95
Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 91.35% Mismatches: 2
Query Match: 36.04% Indels: 0
DB: 6 Gaps: 0

US-10-017-410-2 (1-275) x CO736906 (1-315)
Qy 172 ValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeuPheCys 191
:::
Db 1 ATGGGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCTGTTGGACCTGGCCCTGTTCTGC 60
|||||
Qy 192 TrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPheProTyTrLeuHis 211
|||||
Db 61 TGGATCAGTGACCGAGCTTTCTGGAGCTGTGTCATCTTCAACTTCCCTTACCTGCAC 120
|||||
Qy 212 CysValTrpHisIleLeuLeuAlaSerTyTrLeuGlyCysValCysPheAlaTyTr 231
|||||
Db 121 TGCATGTGGCAGCATCTCTCATCTGCTTGTCTGCTTACCTGGGCTGTGTATGCTTTCCTAC 180
|||||
Qy 232 PheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPheTrpProSerGlu 251
|||||
Db 181 TTTGATGCTGCTCAGAGATTCTCTGAGCAAGGCCCTGTCTATCAAAATTCTGGCCCGAGGAG 240
|||||
Qy 252 LysTrpAlaPheIleGlyValProTyTrValSerLeuLeuCysAlaHisLysLysSerPro 271
|||||
Db 241 AAATGGGCTTCAATGTGGTGTCCCTATGTGTCCCTCTGTGTCCTCTGTGTCACCAAGAAATCATCA 300
|||||
Qy 272 ValLysIleThr 275
|||||

```


AUTHORS

Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof

JOURNAL

Patent: WO 02068579-A 22840 06-SEP-2002;

FEATURES

PE Corporation (NY) (US)
Location/Qualifiers
source
1. .315
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4,45e-43 Length: 315
Score: 545.00 Matches: 95
Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 91.35% Mismatches: 2
Query Match: 36.04% Indels: 0
DB: 6 Gaps: 0

US-10-017-410-2 (1-275) x CQ736906 (1-315)

QY 172 ValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpThrLeuAlaLeuPheCys 191
Db 1 ATCGGTGTGTTAAGCTGGGCTCTTCTCGGCTCTGGTGGACCTGGGCCCTGTTCTGC 60
QY 192 TrpLysSerAspGlnAlaPheCysGluLeuSerSerPheHisPheProTyrLeuHis 211
Db 61 TGGATCAGTGACCGGCTTCTCGGAGCTGCTCATCTCCATCAATCTCCCTACCTGCAC 120
QY 212 CysValTrpHisIleLeuLysLeuAlaSerTyrLeuGlyCysValCysPheAlaTyr 231
Db 121 TGCATGTGCACATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 232 PheAspAlaSerGluLeuProGluGlnGlyProValIleArgPheTrpProSerGlu 251
Db 181 TTTGATGCTGCTCAGAGATTCTCGAGCAAGGCTGTCATCAATCTGGCCAGGAG 240
QY 252 LysTrpAlaPheIleGlyValProTyrValSerLeuLeuCysAlaHisLysLysSerPro 271
Db 241 AATAGGCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 272 ValLysIleThr 275
Db 301 GTCAGACCAACG 312

RESULT 15

AF347023 822 bp mRNA linear ROD 11-AUG-2003
LOCUS Mus musculus alkaline ceramidase mRNA, complete cds.
DEFINITION AF347023
ACCESSION AF347023
VERSION AF347023.1 GI:19070364
KEYWORDS Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 822)

REFERENCE

1 (bases 1 to 822)
Mao, C., Xu, R., Szulc, Z.M., Bielawski, J., Becker, K.P., Bielawska, A.,
Galadari, S.H., Hu, W. and Obeid, L.M.
Cloning and Characterization of a Mouse Endoplasmic Reticulum
Alkaline Ceramidase: AN ENZYME THAT PREFERENTIALLY REGULATES
METABOLISM OF VERY LONG CHAIN CERAMIDES

JOURNAL

J. Biol. Chem. 278 (33), 31184-31191 (2003)

PUBMED

12783875

REFERENCE

2 (bases 1 to 822)
Mao, C., Xu, R. and Obeid, L.M.
Molecular cloning and characterization of a novel murine alkaline
ceramidase

AUTHORS

1 (bases 1 to 822)
Mao, C., Xu, R. and Obeid, L.M.

TITLE

1 (bases 1 to 822)
Mao, C., Xu, R. and Obeid, L.M.

REFERENCE

3 (bases 1 to 822)
Mao, C., Xu, R. and Obeid, L.M.

TITLE

JOURNAL
Submitted (08-FEB-2001) Medicine, Medical University of South
Carolina, 171 Ashley Avenue, Charleston, SC 29425, USA

FEATURES

Location/Qualifiers
source
1. .822
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

CDS

1. .822
/function="deacylates ceramide to generate sphingosine and
a fatty acid"
/codon_start=1
/product="alkaline ceramidase"
/protein_id="AAL83821.1"
/db_xref="GI:19070365"
/translation="MHVPGTRAMSSIFAYQSEVDWCSNFOHSELVAFYNTFNSV
FFLIQGLMMFLMFLPYAQRFCYGVSLFMLIGLFSYFHTLFLGLLEISIL
WLAGSYVLPKPCYFKEVKGKRGVYSLVITITISFLTFKVTWVAYALNSAI
HILYIVRTEYKKIRDDLRHLIAVSVLWAAALTSWISDRVLCSPQRIHFYHLSW
HVLISITPPYIVTALVDKVEMPDKTLKVHWPDRSDWVIGLPIVEIQENDKNC"

ORIGIN

Alignment Scores:
Pred. No.: 5,68e-40 Length: 822
Score: 517.50 Matches: 101
Percent Similarity: 57.09% Conservative: 44
Best Local Similarity: 39.76% Mismatches: 100
Query Match: 34.23% Indels: 9
DB: 10 Gaps: 3

US-10-017-410-2 (1-275) x AF347023 (1-822)

QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 52 AGTTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 111
QY 34 TyrAsnThrIleSerAsnValLeuPheIleLeuProIleCysMetCysLeuPhe 53
Db 112 TACATACGTTGAGCAATGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 171
QY 54 ArgGlnTyrAla-----ThrCysPheAsnSerGlyIleTyrLeuIleTrpThr 69
Db 172 CATCCGTATGCCAGAACGCTACCGGTGTTTC-----TATGAGTGTGAGTC 219
QY 70 LeuLeuValValGlyIleGlySerValTyrPheHisAlaThrIleSerPheLeuGly 89
Db 220 CTCCTTCATGCTCATGGTCTGTCTTCATGATTTCCACATGACATCAGCTCAGCTTCTGGGA 279
QY 90 GlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPhe 109
Db 280 CAGCTGCTGATGAGATCTCCATCTCTGTTGTTGGCCAGTGGATGATGATGATGATGATG 339
QY 110 ProArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaValVal 129
Db 340 CCCCCTTGCTATTTTCCCAAGTTCGTCAAGGGGACACAGGTTTCTACTTCAGCTGCTGGA 399
QY 130 CysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsnIle 149
Db 400 ACTATAACCATATTATATACAGCACCTTCTTGACGCTTCGTGAAGCCCACTCATGATCAT 459
QY 150 SerLeuMetIleLeuGlyLeuProCysThrAlaLeuValAlaGluLeuLysArgCys 169
Db 460 GCTCTCAACAGCATCGCCATCCACATCTCTATCATTTGTCGCACAGAGTACAAGAAGATT 519
QY 170 AspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpThrIleAlaLeu 189
Db 520 AGGGATGATGATCTTCGGCATCTGATGCGGTTTCTGTGTTCTTATGCGCCGCTGACATG 579
QY 190 PheCysTrpIleSerAspGlnAlaPheCysGluLeuSerSerPheHisPheProTyr 209
Db 580 ACCAGCTGGATCAGTACCGGTACTTTCAGCTTCTGCGAGGGATTCCTTCTACTACTAC 639
QY 210 LeuHisCysValTrpHisIleLeuIleCysLeuAlaSerTyrLeuGlyCysValCysPhe 229

```

        |||||   :::||| ||| ||| |||   :::   ||| |||
Db 640 CTGCACAGCATTTGGCACGTCCTCATAGCATCACATTTCCTTATGGTATCGTGACCATG 699
Qy 230 AlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPheTyrPro 249
        |||   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 GCCCTGGTGATGCAAAGTATGAGATGCCAGATATAAACCCCTCAAAGTCCACTACTGGCCC 759
Qy 250 SerGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263
        :::   |||   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 CGGACAGCTGGGTC---ATCGGGCTACCCCTATGTGGAGATC 798

```

Search completed: November 21, 2004, 08:54:39
Job time : 3506 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 20, 2004, 21:28:52 ; Search time 433 Seconds

(without alignments)
3333.928 Million cell updates/sec

Title: US-10-017-410-2

Perfect score: 1512
Sequence: 1 MGAPHWDLHRAGSEVDWC.....IGVPYVSLCAHKKSPVKIT 275

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10017410/runat_19112004_133619_18988/app.query.fasta_1.455
-DB-N Geneseq_23Sep04 -CPM=fastap -SUFFIX=ring -MINMATCH=0 -1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410 @CGN 1 1 708 @runat_19112004_133619_18988 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database : N Geneseq_23Sep04.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1512	100.0	4175	6 ABS54148	Ab854148 Mouse cDN
2	1433	94.8	4202	8 ACC50961	Acc50961 Human bla
3	1433	94.8	4212	11 ADN38833	Adn38833 Cancer/an
4	1404	92.9	828	6 ABS54149	Ab54149 Human cDN
5	1373.5	90.8	822	4 AAH48641	Aah48641 Human cer
6	1110.5	73.4	1527	10 ADL06640	Adl06640 Human 3T3

ALIGNMENTS

RESULT 1
ABS54148
ID ABS54148 standard; cDNA; 4175 BP.
XX
AC ABS54148;
XX
DT 25-NOV-2002 (first entry)
XX
DE Mouse cDNA encoding a liver tumour marker protein, CRG-L1.
XX
KW Mouse; ss; gene; liver cancer; liver tumour; CRG-L1;
KW hepatocellular cancer.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 35..862
FT /tag=a
FT /product= "Liver tumour marker protein, CRG-L1"
XX
PN US2002115094-A1.
XX
PD 22-AUG-2002.
XX
PF 14-DEC-2001; 2001US-00017410.
XX
PR 14-DEC-2000; 2000US-0255674P.
XX
PA (FARN/) FARNHAM P J.
PA (GRAV/) GRAVEEL C R.

7	1110.5	73.4	1527	10	ADL06642
8	1087	71.9	1623	10	ADL06642
9	584	38.6	487	10	ADI21948
10	568	37.6	1215	4	AAK51566
11	568	37.6	1631	4	AAK52550
12	549.5	36.3	792	4	AAH48639
13	455	30.1	537	4	ABL24043
14	320	21.2	4758	4	ABL24042
15	264.5	17.5	4071	4	ABL30432
16	209.5	13.9	801	4	AAH48640
17	209.5	13.9	1063	6	ABZ11993
18	209.5	13.9	1063	12	ADM44511
19	208.5	13.8	2744	3	AAAG6499
20	199	13.2	1239	3	AAC40178
21	186	12.3	35425	10	ADC87616
22	182	12.0	2058	4	AAH15072
23	182	12.0	4646	5	ABV27900
24	182	12.0	4646	5	ABV28037
25	182	12.0	4646	5	ABV22202
26	182	12.0	4646	5	ABV22064
27	161	10.6	636	6	ABL90053
28	159	10.5	639	5	AAAS79600
29	143.5	9.5	1148	3	AAF13013
30	142	9.4	748	4	AAH07955
31	136	9.0	354	10	ABX74425
32	123.5	8.2	1498	10	ADC87192
33	118.5	7.8	6755	2	AAV21511
34	116.5	7.7	5325	4	ABL23204
35	113	7.5	2842	4	ABL23205
36	112	7.4	1501	10	ADC87290
37	111	7.3	2122	5	AAAS80960
38	111	7.3	2521	5	AAAS66068
39	111	7.3	50335	9	AAAS8280
40	111	7.3	226475	9	AAAS8279
41	110.5	7.3	349981	10	ADC87619
42	110.5	7.3	349989	10	ADC86916
43	110	7.3	1147	8	ABZ52533
44	109	7.2	1282	10	ADC87232
45	108.5	7.2	5842	12	ADJ12462

XX Farnham PJ, Graveel CR;
 XX WPI; 2002-706409/76.
 DR P-PSDB; ABG32880.
 XX
 PT Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
 PT liver cancer, is differentially expressed in liver tumors relative to
 PT normal liver tissues.
 XX
 XX Claim 2; Page 4-7; 11pp; English.
 XX The invention relates to a polypeptide designated as CRG-L1, which is
 CC differentially expressed in liver tumors relative to normal expression
 CC in normal liver tissues, designated CRG-L1. Also included are the
 CC encoding polynucleotides (in the case of the human sequence, mapping to
 CC chromosome 9p), expression constructs, host cells, anti-CRG-L1
 CC antibodies, identifying modulators of CRG-L1, and the use of the CRG-L1
 CC sequence in the diagnosis of hepatocellular cancer in tumour cells from a
 CC liver of a human or non-human animal. The CRG-L1 protein and
 CC polynucleotide are useful as diagnostic markers for a liver cancer in
 CC humans and non-human animals, and as a system for assessing putative
 CC therapeutic agents. The present sequence encodes mouse CRG-L1
 XX
 SQ Sequence 4175 BP; 962 A; 971 C; 1012 G; 1230 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,98e-142 Length: 4175
 Score: 1512.00 Matches: 275
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-017-410-2 (1-275) x ABS54148 (1-4175)

QY 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
 DB 35 ATGGGGCCCGCCGACCTGGTGGGACCCACCTGGGGCTGGGAGTTCGGAGGTGGATTGGTGC 94
 QY 21 GluAspAsnTrpThrIleValProAlaIleAlaGluPheTrpAsnThrIleSerAsnVal 40
 DB 95 GAGGACAACTACACTATCGTGGCTGGCCATTCGCCAGTCTTACAAACAGATCAGCAACGTC 154
 QY 41 LeuPhePheIleLeuProIleCysMetCysLeuPheArgGlnTrpAlaThrCysPhe 60
 DB 155 TTGTTTTTCAATTTACTCCCATCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
 QY 61 AsnSerGlyIleTrpLeuIleTrpThrLeuLeuValValValGlyIleGlySerValTyr 80
 DB 215 AACAGCGGCATCTACTTATATGAGCGCTCCTAGTGTAGTGGGATTGGATCTGTCTAC 274
 QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
 DB 275 TTCCATGCAACGGTGGTTCCTGGGTGATGCTTGTGATGCTTGTGATGCTTGTGATGCTT 334
 QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTrpLeuProIlePheArgAsn 120
 DB 335 CTGATGTGTGCTTTGGCCAGTGGTGTCCAGAGGATTTTACCAAGATCTTTCGGAAT 394
 QY 121 AspArgGlyArgPheLeuAlaValValCysValLeuSerAlaIleThrTrpCysLeuAla 140
 DB 395 GACAGGGGCGAGTTCAAGGCGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 454
 QY 141 PheIleLysProAlaIleAsnAlaIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
 DB 455 TTTATCAAGCCCGCCATCAACAATATTTCCCTGATGATTTCTGGGACTTCATGCACTGGG 514
 QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLeuGlyLeuPhe 180
 DB 515 CTGCTGTGTGACAGCTGAAGAGGTGTGCAATGTGCGTGTGTTTAACTGGGCTCTTC 574
 QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGlu 200

DB 575 TCTGGCCCTCGTGGACTCTGGCTCTCTCTGATCAGCAGCAGCAGCAGCAGCAGCAGCAG 634
 QY 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeuIleCysLeu 220
 DB 635 CTGCT 694
 QY 221 AlaSerTyrIleuGlyCysValCysPheAlaTyrPheAspAlaIleSerGluIleProGlu 240
 DB 695 GCTTCGTACCTGGGCTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 754
 QY 241 GlnGlyProValIleArgPheTrpProSerGluIleValTrpAlaPheIleGlyValProTyr 260
 DB 755 CAAGGTCAGTCTATCAGATTCTGGCCAGCAGAAATGGGCTTTATTTGGTGTCCCTTAT 814
 QY 261 ValSerLeuLeuCysAlaHisIleLysSerProValIleTrp 275
 DB 815 GTGTCT 859

RESULT 2
 ACC50961
 ID ACC50961 standard; cDNA; 4202 BP.
 XX
 AC ACC50961;
 XX
 DT 12-JUN-2003 (first entry)
 XX
 DE Human bladder cancer associated cDNA sequence SEQ ID NO:21.
 XX
 KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO2003003906-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 03-JUL-2002; 2002WO-US021338.
 XX
 PR 03-JUL-2001; 2001US-0302814P.
 PR 03-AUG-2001; 2001US-0310099P.
 PR 08-NOV-2001; 2001US-0343705P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (BOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Mack DH, Aziz N;
 XX
 DR WPI; 2003-201532/19.
 DR P-PSDB; ABR48156.
 XX
 PT Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT bladder cancer-associated polynucleotide or antibody.
 XX
 PS Claim 6; Page 235; 307pp; English.
 XX
 CC The present invention describes a method for detecting a bladder cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridizes to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications

XX SQ Sequence 4202 BP; 1045 A; 892 C; 927 G; 1338 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2 9e-134 Length: 4202
Score: 1433.00 Matches: 256
Percent Similarity: 98.18% Conservative: 14
Best Local Similarity: 93.09% Mismatches: 5
Query Match: 94.78% Indels: 0
DB: 8 Gaps: 0

US-10-017-410-2 (1-275) x ACC50961 (1-4202)

Qy 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
Db 77 ATGGGGCCCGCCGACCTGGTGGGACACAGCTGCAGGCTGGTAGCTCGAGGTGGACTGGTGC 136
Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 137 GAGGACAACTACACCATCGTGGCTGCTATCGCGAGTTCACACAGATCAGCAATGTC 196
Qy 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
Db 197 TTATTTTCATTATACGCCCACTGATGTCTGTTCGTGATGATGCAATGCTTC 256
Qy 61 AsnSerGlyIleTyrLeuIleTyrThrLeuLeuValValGlyIleGlySerValTyr 80
Db 257 AACAGTGGCATCTACTTAATCTGACTCTTTTGGTCTAGTGGGAATGGATCGCTCTAC 316
Qy 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTyrVal 100
Db 317 TTCATGCAACCTTATGTTTCTGGGTGATGCTTGTATGAGTTCAGTCTTTCGGTT 376
Qy 101 LeuMetCysAlaLeuAlaMetTyrPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db 377 CTGATGTGCTTGGCCATCTGCTGCCAGAGGTATCTACCAAGATCTTTCGAAT 436
Qy 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThrThrCysLeuAla 140
Db 437 GACCGGGTAGGTTCAGAGTGGTGGTTCAGTGTCTGCTGGTTCAGCGTTCGCTGGCA 496
Qy 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
Db 497 TTTGTCAAGCTGCAATCAACATCTCTGATGACCCCTGGAGTTCCTTGCACGTCA 556
Qy 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
Db 557 CTGCTCATCGCAGAGCTAAGAGGTGTGACACATGTGTGTGTTAAGTGGGCTCTTC 616
Qy 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGlu 200
Db 617 TCGGGCTCTGGTGGACCTTGGCTGCTGCTGGATCAGTGACCGAGCTTTCGCGAG 676
Qy 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTyrHisIleLeuIleCysLeu 220
Db 677 CTGCTGTCATCTTCAACTCTCCCTACTGCACTGATGTCATGTCATCTTCGCTT 736
Qy 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db 737 GCTGCTACCTGGGCTGTGTATGCTTGGCTACTTGTGCTGCTGCTGATGATTCCTGAG 796
Qy 241 GlnGlyProValIleArgPheTrpProSerGluLysTrpAlaPheIleGlyValProTyr 260
Db 797 CAAGGCCCTGTGATCAAGTCTGGGCCCAATGAGAAATGGGCTTTCATTGTGTCCTAT 856
Qy 261 ValSerLeuLeuCysAlaHisLysLysSerProValLysIleThr 275
Db 857 GTGTCCCTCTGTGTGCAACAGAAATCATCAGTCAAGATCAGC 901

RESULT 3

ADN38833

ID ADN38833 standard; cDNA; 4212 BP.

XX

AC ADN38833;

XX DT 17-JUN-2004 (first entry)

XX DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:151.

XX KW Human; differential expression; cancer; angiogenic disorder;
XX KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
XX KW inflammatory disease; autoimmune disease;
XX KW retinal neovascularisation syndrome; scarring; uterine fibroid;
XX KW detection; diagnosis; prognosis; drug screening; drug targeting;
XX KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX KW vulnery; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX XX WO2003042661-A2.

XX XX 22-MAY-2003.

XX XX 13-NOV-2002; 2002WO-US036810.

XX XX 13-NOV-2001; 2001US-0350666P.

XX PR 21-NOV-2001; 2001US-0332464P.

XX PR 29-NOV-2001; 2001US-0334393P.

XX PR 03-DEC-2001; 2001US-0335394P.

XX PR 14-DEC-2001; 2001US-0340376P.

XX PR 08-JAN-2002; 2002US-0347211P.

XX PR 10-JAN-2002; 2002US-0347349P.

XX PR 08-FEB-2002; 2002US-035250P.

XX PR 13-FEB-2002; 2002US-0356714P.

XX PR 20-FEB-2002; 2002US-0359077P.

XX PR 29-MAR-2002; 2002US-0368809P.

XX PR 12-APR-2002; 2002US-0370110P.

XX PR 12-APR-2002; 2002US-037246P.

XX PR 05-JUN-2002; 2002US-0386614P.

XX PR 16-JUL-2002; 2002US-0396839P.

XX PR 22-JUL-2002; 2002US-0397775P.

XX PR 22-JUL-2002; 2002US-0397845P.

XX PR 09-SEP-2002; 2002US-0409450P.

XX XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX XX WPI: 2003-468649/44.

XX DR P-PSDB; ADN38834.

XX XX

XX XX Determining the presence or absence of a pathological cell in a patient,
XX XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX XX a nucleic acid in a biological sample.

XX XX Claim 8; SEQ ID NO 151; 1385pp; English.

XX XX The invention relates to nucleic acids and proteins (ADN38833-ADN40064)
XX XX whose expression is upregulated or downregulated in specific cancers or
XX XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX XX of determining the presence or absence of a pathological cell in a
XX XX patient by detecting a nucleic acid at least 80% identical to those of
XX XX the invention or by detecting a polypeptide of the invention. The
XX XX invention also relates to expression vectors and host cells comprising a
XX XX nucleic acid of the invention; antibodies which specifically bind a
XX XX polypeptide of the invention; use of such antibodies for drug targeting;
XX XX and methods of screening for modulators of activity or expression of the
XX XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX XX antibodies and methods are useful for diagnosing, prognosing and treating
XX XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX XX neovascularisation syndromes, scarring and uterine fibroids. They may
XX XX also be useful in wound healing and in contraception. The present
XX XX sequence represents a nucleic acid sequence of the invention.

XX

SQ Sequence 4212 BP; 1047 A; 894 C; 932 G; 1339 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,91e-134 Length: 4212
Score: 1433.00 Matches: 256
Percent Similarity: 98.18% Conservative: 14
Best Local Similarity: 93.09% Mismatches: 5
Query Match: 94.78% Indels: 0
DB: 11 Gaps: 0

US-10-017-410-2 (1-275) x ADN38833 (1-4212)

```
QY 1 MetGlyAlaProHisTrpPheAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
DB 87 ATGGGCGCCCGACCTGGTGGGACCGAGCTGAGCTGGTAGCTCGGAGGTGGACTGGTGC 146
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
DB 147 GAGGACAACTACACCATCGTGCCTGCTATCGCGAGTCTTACAAACAGATCAGCAATGTC 206
QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
DB 207 TTATTTTTCATTTTACCGCCCATCTGCATGTGCTTTCGTCAGTATGCAACATGCTTC 266
QY 61 AsnSerGlyIleTyrLeuIleTyrThrLeuLeuValValGlyIleGlySerValTyr 80
DB 267 AACAGTGGCATCTACTTAATCTGGACTCTTTGGTGTAGTGGAAATGGATCCGCTAC 326
QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
DB 327 TTCCATGCAACCTTAGTTCTTGGGTGAGTGTGATGTAACCTTCAGTCCCTTTGGTT 386
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgTyrIleuProLysIlePheArgAsn 120
DB 387 CTGATGTGTCTTGGCCATCTGGTTCCCAAGAGTATCTACCAAGATCTTTTCGAAT 446
QY 121 AspArgGlyArgPheLeuAlaValValCysValLeuSerAlaIleThrThrCysLeuAla 140
DB 447 GACCGGGTAGGTTCAAGTGGTGGTCAAGTCTCTGCGTACGAGCGCTGGCA 506
QY 141 PheIleLysProAlaIleAsnAniIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
DB 507 TTTGTCAAGCCTGCCATCAACAACATCTCTGTATGACCTGGAGTTCCTTGCACTGCA 566
QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
DB 567 CTGCTCATCGCAGAGTAAAGAGGTGTGACACATGCGTGTGTTAAAGCTGGGCTCTTC 626
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGlu 200
DB 627 TGGGGCTCTGGTGGACCTGGCCCTGTTCTGCTGATCAGTACCGAGCTTCTGCGAG 686
QY 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeuIleCysLeu 220
DB 687 CTGCTGTATCTCTCAACTCCCTACCTGCACTGATGGCAGCATCTCATCTGCTT 746
QY 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
DB 747 GCTGCTACTCGGCTGTGTATGCTTGTGCTACTTTGATGCTGCCTCAGAGATTCCTGAG 806
QY 241 GlnGlyProValIleArgPheTrpProSerGluLysTrpAlaPheIleGlyValProTyr 260
DB 807 CAAGGCCCTGTATCAAGTTCTGGCCCAATGAGAAATGGCCCTTCATTTGGTGTCCCTAT 866
QY 261 ValSerLeuLeuCysAlaHisLysLysSerProValLysIleThr 275
DB 867 GTGTGCTCTCTGTGTGCCAACAGAAATCATCAGTCAAGATCAAG 911
```

RESULT 4

ABS54149

ID ABS54149 standard; cDNA; 828 BP.

XX

AC ABS54149;

```
XX 25-NOV-2002 (first entry)
DT
XX
XX Human cDNA encoding a liver tumour marker protein, CRG-L1.
XX
XX Human; ss; gene; liver cancer; liver tumour; CRG-L1;
XX hepatocellular cancer; chromosome 9p.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..828
XX FT /tag= a
XX FT /product= "Liver tumour marker protein, CRG-L1"
XX
XX US2002115094-A1.
XX
XX 22-AUG-2002.
XX
XX 14-DEC-2001; 2001US-00017410.
XX
XX 14-DEC-2000; 2000US-0255674P.
XX
XX (FARN/) FARNHAM P J.
XX (GRAV/) GRAVEEL C R.
XX
XX Farnham PJ, Graveel CR;
XX
XX WPI; 2002-706409/76.
XX P-PSDB; ABG32881.
XX
XX Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
XX liver cancer, is differentially expressed in liver tumors relative to
XX normal liver tissues.
XX
XX Claim 2; Page 8; 11pp; English.
XX
XX The invention relates to a polypeptide designated as CRG-L1, which is
XX differentially expressed in liver tumours relative to normal expression
XX in normal liver tissues, designated CRG-L1. Also included are the
XX encoding polynucleotides (in the case of the human sequence, mapping to
XX chromosome 9p), expression constructs, host cells, anti-CRG-L1
XX antibodies, identifying modulators of CRG-L1, and the use of the CRG-L1
XX sequence in the diagnosis of hepatocellular cancer in tumour cells from a
XX liver of a human or non-human animal. The CRG-L1 protein and
XX polynucleotide are useful as diagnostic markers for a liver cancer in
XX humans and non-human animals, and as a system for assessing putative
XX therapeutic agents. The present sequence encodes human CRG-L1
XX
XX Sequence 828 BP; 157 A; 216 C; 211 G; 244 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.: 2,96e-132 Length: 828
Score: 1404.00 Matches: 251
Percent Similarity: 96.73% Conservative: 15
Best Local Similarity: 91.27% Mismatches: 9
Query Match: 92.86% Indels: 0
DB: 6 Gaps: 0

US-10-017-410-2 (1-275) x ABS54149 (1-828)

```
QY 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
DB 1 ATGGGCGCCCGACCTGGTGGGACCGAGCTGAGCTGGTAGCTCGGAGGTGGACTGGTGC 60
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
DB 61 GAGGACAACTACACCATCGTGCCTGCTATCGCGAGTCTTACAAACAGATCAGCAATGTC 120
QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
DB 121 TTATTTTTCATTTTACCGCCCATCTGCATGTGCTTGTGTTGATGAGTATGCAACATGCTTC 180
```


Qy 263 LeuLeuCysAlaHisLysLysSerProValLysIleThr 275
 Db 1199 CTCCTGTGTGCAACAAAGAAATCATCATGTCAGTCAAGATCAG 1237

RESULT 7

ADL06642
 ID ADL06642 standard; DNA; 1527 BP.

XX AC ADL06642;

XX XX 06-MAY-2004 (first entry)

XX XX Human 3T3 cell conversion promoter PP11646 DNA.

XX XX 3T3 cell conversion; promoter; human; gene; ds.

XX XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 218..640

XX FT /*tag= a

XX FT /product= "PP11646"

XX CN1403477-A.

XX PD 19-MAR-2003.

XX PF 12-SEP-2001; 2001CN-00126725.

XX PR 12-SEP-2001; 2001CN-00126725.

XX (SHAN-) SHANGHAI XINSHIJI GENE TECHN DEV CO LTD.

XX Gu J, Yang S;

XX WPI; 2003-494226/47.

XX P-PSDB; ADL06641.

XX Human protein with function of promoting 3T3 cell conversion and its coding sequence.
 PS Example 1; SEQ ID NO 3; 41pp; Chinese.
 CC This invention describes a novel human protein with 3T3 cell conversion promoting function, polynucleotides encoding the polypeptide and the recombinant process of producing the polypeptide. The present invention also discloses the agonist resisting the polypeptide and its treatment effect. The present invention also discloses the application of the polynucleotides encoding the human protein with 3T3 cell conversion promoting function.

XX SQ Sequence 1527 BP; 376 A; 357 C; 383 G; 411 T; 0 U; 0 Other;

Alignment Scores:

Fred. No.: 3.05e-102 Length: 1527
 Score: 1110.50 Matches: 222
 Percent Similarity: 66.20% Conservative: 13
 Best Local Similarity: 62.54% Mismatches: 4
 Query Match: 73.45% Indels: 116
 DB: 10 Gaps: 2

US-10-017-410-2 (1-275) x ADL06642 (1-1527)

Qy 37 IleSerAsnValLeuPheIleLeuProPheLeuPheArgGlnTyr 56

Db 179 ATCAGCAATGCTTATTTTCATTTTACGCCCATCTGCATGTGTTCTTCAGTAT 238

Qy 57 AlaThrCysPheAsnSerGlyIleTyrLeuIleThrLeuValValGlyIle 76

Db 239 GCAACATGCTTCAACAGTGGCATCTACTTAATCTGCACCTCTTTGTGTAGTGGAAAT 298

Qy 77 GlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGlnLeuAla 96

Db 299 GGATCCGCTACTTCCATGCAACCCCTAGTTTCTTGGTGCAGATGCTTGATGAACCTTGC 358
 Qy 97 IleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLys 116
 Db 359 GTCCTTTGGGTTCTGATGTGTGCTTTGGCCATGTGTTCGCCAAGAGGTATCTACCAAG 418
 Qy 117 IlePheArgAsnAspArgGlyArgPheIysAlaValValCysValLeuSerAlaIleThr 136
 Db 419 ATCTTTCCGAATGACCCGGGTAGTTCAGAGTGTGTGTCAGTCTCTGTCGCGGTACG 478
 Qy 137 ThrCysLeuAlaPheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeu 156
 Db 479 ACGTGCTTGGCATTTGTCAAGCCTGCCATCAACACATCTCTCTGATGACCTTGGAGTT 538
 Qy 157 ProCysThrAlaLeuLeuValAlaGluLeuLys----- 167
 Db 539 CTTTGCACCTGCACTGCTCATCGCAGAGCTAAAGAGGCATGAGAGAACCCAGCAAGGAGA 598
 Qy 167 ----- 167
 Db 599 CACAGGAAAGTGGCCAGCAAGAGGTGGAGCAAGGTCTGACGATGATGACTCTCTGG 658
 Qy 168 -----ArgCysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTr 184
 Db 659 ACCCCCGTGCAAGTGTGACACATGCGTGTGTTTAAAGCTGGCCCTCTTCTCGGGCTCTG 718
 Qy 184 pTrpThrIleuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSerSe 204
 Db 719 GTGACCCCTGGCCCTGTTCTGTCGATCAGTGACCGAGCTTTCGCGAGCTGCTGTCATC 778
 Qy 204 rPheHisPheProTyrLeuHisCysValTrp----- 214
 Db 779 CTTCAACTTCCCTTACCTGCATGCTGTGTGGACGCTCAGATGTTTCAGAACCCACTG 838
 Qy 214 ----- 214
 Db 839 AAAGCACACTTTGTAAAGAGCTGAATTGACTCCAGGAGCAGGCTTAGCCGGAACGAAG 898
 Qy 214 ----- 214
 Db 899 GCCTGATTGTCATGCTTAGCAGGACTATTTAAGAACATTTCTAAACGGGTCCCAAAATGC 958
 Qy 214 ----- 214
 Db 959 TGAAGATTATCATCTCTGCTATTCAGGAGTGTGTGTAATCTTCATCCTTCTCTAAATA 1018
 Qy 215 -----HisIleLeuIleCysLeuAlaSer 222
 Db 1019 GAAGAGGCCCTGGGCTGCAACTACAGTCAGTCAGCAAGGCACATCTCATCTGCTTGTGCC 1078
 Qy 223 TyrLeuGlyCysValCysPheAlaTyrPheAspAlaIleSerGluIleProGluGlnGly 242
 Db 1079 TACCTGGGCTGTGTATGCTTTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1138
 Qy 243 ProValIleArgPheTrpProSerGluLysTrpAlaPheIleGlyValProTyrValSer 262
 Db 1139 CTTGTCTATCAAGTCTTCTGCCCAATGAGAAATGGGCTTTCATGTGGTGTCCCCATGTGTGCC 1198
 Qy 263 LeuLeuCysAlaHisLysLysSerProValLysIleThr 275
 Db 1199 CTCCTGTGTGCCAACAGAAATCATCATGATCAGTCAAGATCAG 1237

RESULT 8

AD121948

ID AD121948 standard; cDNA; 1623 BP.

XX AC AD121948;

XX DT 15-APR-2004 (first entry)

XX DE Novel human protein cDNA #207.

XX KW forensic; nutritional source; damaged tissue; diseased tissue;

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
PI Haley-Vicente D;
XX
XX
DR WPI: 2003-354603/33.
DR P-PSDB; ADI21688.
XX
XX New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.
XX
XX Example 2; SEQ ID NO 719; 156pp; English.
XX
XX The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers,
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents a novel human expressed sequence tag, EST.
XX
SQ Sequence 487 BP; 81 A; 138 C; 131 G; 134 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 1.12e-49 Length: 487
Score: 584.00 Matches: 105
Percent Similarity: 98.17% Conservative: 2
Best Local Similarity: 96.33% Mismatches: 2
Query Match: 38.62% Indels: 0
DB: 10 Gaps: 0
US-10-017-410-2 (1-275) x ADI21468 (1-487)
QY 1 MetGlyAlaProHisTyrTrpAspHisLeuArgAlaGlySerSerGluValAspTyrCys 20
Db 159 ATGGCGCGCCCGCACTGGTGGGACCGAGCTGCAGGCTGGTAGCTCGAGGTGGACTGGTGC 218
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 219 GAGGACAACTACACATCGTGCCTGTATCCCGAGTTCTTCAACACCATCAGCAATGTC 278
QY 41 LeuPhePheIleLeuProPheIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
Db 279 TTAATTTTCATTTTACCGCCCATCGCATGTGCTGTGTTTCTGTCAGTATGCAATGCTTC 338
QY 61 AsnSerGlyIleTyrLeuIleTyrThrIleValValValGlyIleGlySerValTyr 80
Db 339 AACAGTGGCATCTACTTAATCTGGACTCTTTTGGTGTGTAGTGGGAATTCGATCCGCTCAC 398
QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTyrVal 100
Db 399 TTCATGCAACCCCTAGTTTCTTGGGTGAGATGCTTGATGAACCTTGCAGTCCCTTGGGT 458
QY 101 LeuMetCysAlaLeuAlaMetTyrPhe 109
Db 459 CTGATGTGTGCTTTAGCCATGTGGTTC 485

RESULT 10
ID AAK51566
AAK51566 standard; cDNA; 1215 BP.
XX
AC AAK51566;
XX

DT 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 111.
DE
XX
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
OS Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 30-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR P-PSDB; AAM78433.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 1; Page 750-751; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 1215 BP; 213 A; 374 C; 309 G; 319 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.54e-47 Length: 1215
Score: 568.00 Matches: 112
Percent Similarity: 53.97% Conservative: 51
Best Local Similarity: 37.09% Mismatches: 108
Query Match: 37.57% Indels: 31
DB: 4 Gaps: 3
US-10-017-410-2 (1-275) x AAK51566 (1-1215)
QY 2 GlyAlaProHisTyrTrp-----AspHisLeuArgAlaGly-Se 14
Db 22 GCGCAGCGCGCTGGTGGTGGCCACTGAGACAAGATGCTTCGCTATCAGAG 81
QY 14 rSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAlaGluPheTy 34
Db 82 CTCGAGGTGGACTGGTGTGTGAGAGCACTTCAGTACTCGGAGCTGGTGGCGAGTTCTA 141

QY 34 rAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPheAr 54
 Db 142 CAACACGTTCTCCAAATATCCCTTCTTCATCTTCGGCCACTGATGATGCTCTCTGATGCA 201
 QY 54 gGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThrLeuValValVa 74
 Db 202 CCCGATATGCCAGAGCGCTCCCGCTACATTTAGTTGCTGGGCTCTCTTCATGATCAT 261
 QY 74 lGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAepGl 94
 Db 262 AGCCCTGTTCTCCATGATTTCCACATGACGCTCAGCTTCCTGGCCAGCTGCTGGACGA 321
 QY 94 uLeuAlaIleLeuTyrValLeuMetCysAlaLeuAlaMetTyrPheProArgArgTyrLe 114
 Db 322 GATCGCATCTCTGGCTCTGGCAGTGGCTATAGCATATGATGATGCTGGCTGCTATTT 381
 QY 114 uProLysIlePheArgAsnAspArgGlyArgPheLysAlaValValCysValLeuSerAl 134
 Db 382 CCCCTCTCTTGGGGGAACAGGTCCAGTTTCATCCGCTGCTCTTCATCACCAGCTGT 441
 QY 134 alleThrThrCysLeuAlaPheIleLysProAlaIleAsnAsnIleSerLeuMetIleLe 154
 Db 442 GTCAGCACCTTCTGCTCTCTGGCGCCCGCTCAGCTCAGCGCTTACCGCTCAACAGCAT 501
 QY 154 uGlyLeuProCysThrAlaLeuValAlaGluLeuLysArgCysAspAsnValArgVa 174
 Db 502 TGCCTGTCATCTCTACATGCTGTCAGGAGTACAGGAGACCAAGCAATAGGAGCT 561
 QY 174 lPheLysLeuGlyLeuPheSerGlyLeuTyrThrLeuAlaLeuPheCysTyrIleSe 194
 Db 562 TCGGCACCTGATGAGTCTCCGTGGTGTATGGCGTGTGCTCTGACAGCTGGATCAG 621
 QY 194 rAspGlnAlaPheCysGluLeuLeuSerSerPheHisPheProTyrLeuHisCysValTr 214
 Db 622 TGACCGTCTGCTTTCGACGCTTCTGGCAGAGGATTCATTTCTTCTATCTGCACAGCATCTG 681
 QY 214 pHisIleLeuIleCysLeuAlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAl 234
 Db 682 GCATGCTCATCAGATCATCTCCCTTATGGATGCTGCTGATGCTGCTGGTGGATGC 741
 QY 234 aAlaSerGluIleProGluGlnGlyProValIleArgPheTyrProSerGlyLysTyrPal 254
 Db 742 CAACTATGATGATGCCAGGTGAACCCCTCAAGTCCGCTACTGCTGCGGACAGTTGGCC 801
 QY 254 aPheIleGlyValProTyrValSerLeu----- 263
 Db 802 C---GTGGGGCTGCCCTACGTGGAAATCCGGGGTGATGCAAGGACTGCTGACAGCTGCC 858
 QY 264 -----LeuCysAlaHisLysLysSerProVa 272
 Db 859 AGCCTCTTGACTATCCACACCCCAACACTTGCCTGTGCTTGTGAGAAGATAGCCCGGT 918
 QY 272 lLys 273
 Db 919 CAGG 922

RESULT 11

AAK52550
 ID AAK52550 standard; cDNA; 1631 BP.
 XX
 AC
 XX
 XX
 DT
 DT
 XX
 DE
 XX
 KW
 KW
 KW
 KW
 XX
 XX
 OS

Alignment Scores:
 Pred. No.: 2,26e-47 Length: 1631
 Score: 568.00 Matches: 112
 Percent Similarity: 53.97% Conservative: 51
 Best Local Similarity: 37.09% Mismatches: 108
 Query Match: 37.57% Indels: 31
 DB: 4 Gaps: 3

US-10-017-410-2 (1-275) x AAK52550 (1-1631)

QY 2 GlyAlaProHisTyrTrp-----AspHisLeuArgAlaGly-Se 14
 Db 22 GCGCGACGCGGCTGGTGGTGGCCACTGAGCAAGATGCTAGCATCTCGCTATCAGAG 81
 QY 14 rSerGluValAspTyrCysGluAspAsnTyrThrIleValProAlaIleAlaGluPheTy 34
 Db 82 CTCGAGGAGGAGTGGTGTGAGAGCAACTTCAGTACTCGGAGCTGGTGGCGAGTCTCA 141
 QY 34 rAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPheAr 54
 Db 142 CAACAGTTCTCAATATATCCCTCTCTTCATCTTCGGCCACTGATGATGCTCTCTGATGCA 201
 QY 54 gGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThrLeuValValVa 74
 Db 202 CCCGATATGCCAGAGCGCTCCCGCTACATTTACGTTGTCTGGCTCTCTTCTATGATCAT 261
 QY 74 lGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAepGl 94

XX WO200157190-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US004098.
 XX 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue Au, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 DR P-PSDB; AAM79417.
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 in diagnosis and gene therapy.
 PT Claim 1; Page 4462; 6221pp; English.
 PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC treatment/inhibit cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX Sequence 1631 BP; 247 A; 399 C; 344 G; 350 T; 0 U; 291 Other;

QY 188 aLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuSerSerPheHisPhePr 208
 Db 549 TCGACACGTCGATCAGTACCGCTCGCTTTCAGCTTCGCGAGGAGTTCATTCTT 608
 QY 208 oTyLeuHisCysValTrpHisIleLeuIleCysLeuAlaSerTyrrLeuGlyCysValCy 228
 Db 609 CTATCTGCACAGCATCTGGCATGGCTCATCATCATCATCACCTTCCTTATGGCATGGTAC 668
 QY 228 sPheAlaTyrrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleAlaPheTr 248
 Db 669 CATGGCCTTGGTGGATGCCAACTATGATGATGATGATGATGATGATGATGATGATGATG 728
 QY 248 pProSerGluIleTyrrAlaPheIleGlyValProTyrrValSerLeu 263
 Db 729 GCCTCGGACAGTGGCCCC---GTGGGGCTGGCCCTACGTGGAAATC 771

RESULT 13
 ABL24043
 ID ABL24043 standard; DNA; 537 BP.
 AC ABL24043;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23602.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PR 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 23602; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
 CC AB57072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 537 BP; 102 A; 146 C; 148 G; 141 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.39e-36	Length:	537
Score:	455.00	Matches:	82
Percent Similarity:	69.14%	Conservative:	30
Best Local Similarity:	50.62%	Mismatches:	50
Query Match:	30.09%	Indels:	0
DB:	4	Gaps:	0

US-10-017-410-2 (1-275) x ABL24043 (1-537)

QY 7 TrpAspHisLeuArgAlaGlySerSerGluValAspTrpCysGluAspAsnTyrrThrIle 26
 Db 49 TGGGAGACCTAAGACCCGAGAGCTGCCGTCGACTGGTGGAGGACNACTACTTGTATT 108
 QY 27 ValProAlaIleAlaGluPheTyrrAsnThrIleSerAsnValLeuPhePheIleLeuPro 46
 Db 109 TCGTCCAAATCGCCGAGTTCGTGAACACGTTTAGCAACTTCCTGTTCATCTACTGCG 168
 QY 47 ProIleCysMetCysLeuPheArgGlnTyrrAlaThrCysPheAsnSerGlyIleTyrrLeu 66
 Db 169 CCGTCTCTATATATGCTCTTCAAGGAGTACGGACGCTTTGTGACGCCCGGAATCCACGTC 228
 QY 67 IleTrpThrLeuLeuValValGlyIleGlySerValTyrrPheHisAlaThrLeuSer 86
 Db 229 ATCTGGGTGCTGCTCATCGTGGTGGCTGAGTTCGATGTACTTCCATGCCACTTTGAGT 288
 QY 87 PheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAla 106
 Db 289 CTGATTGGCCAGCTGCTGGACGAACCTGGCCATCTCTGGGTCTTTCATGGCGGCTTTTCG 348
 QY 107 MetTrpPheProArgTyrrLeuProLeuIlePheArgAsnAspArgGlyArgPheLys 126
 Db 349 CTCTTCTATCCGAAGCGATATATCCCAAGTTCGTGAAAACGATCGCAAAACCTTCAGT 408
 QY 127 AlaValValCysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIle 146
 Db 409 TGGCTCATGCTCTTGTGGCGGATTGCTGGACGGGCTTGTGCGTGGTGAAGCCATTGTT 468
 QY 147 AsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeu 166
 Db 469 AACGCCCTTTGTTCTCATGTTTCATGAGTGTGCCGACCATGATGTAATGCTCTACACAGAGCTG 528
 QY 167 LysArg 168
 Db 529 CAGAGG 534

RESULT 14
 ABL24042
 ID ABL24042 standard; DNA; 4758 BP.
 AC ABL24042;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23599.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PR 27-SEP-2001.
 XX
 PR 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX

PS Claim 1; SEQ ID NO 23599; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 4758 BP; 1304 A; 981 C; 980 G; 1493 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.07e-21	Length:	4758
Score:	320.00	Matches:	64
Percent Similarity:	60.00%	Conservative:	29
Best Local Similarity:	41.29%	Mismatches:	38
Query Match:	21.16%	Indels:	24
DB:	4	Gaps:	2

US-10-017-410-2 (1-275) x ABL24042 (1-4758)

Qy 38 SerAsnValLeuPhePheLeuProProLeuCysMetCysLeuPheArgGlnTyrAla 57

Db 3291 AGCAACTTCCTGTTTCACCTACTGCGCCGCTCTTATATGCTCTTCAAGGAGTACGGA 3350

Qy 58 ThrCysPheAsnSerGlyIleTyrLeuLeuTrpHrLeuValValGlyIleGly 77

Db 3351 CGCTTTGTGAGCGCCGGAATCCAGTCATCTGGTGTCTGCTCATCGTGGTGGCTGAGT 3410

Qy 78 SerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIle 97

Db 3411 TCGATGACTTCATGCGACCTTTGAGTCTGATTGGCCAGCTGCTGGAGCACTGGCCATA 3470

Qy 98 LeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIle 117

Db 3471 CTCTGGGTCTTCATGGGGCGCTTTTCGCTCTCTTATCCGAAGCATACTATCCCAAGTTC 3530

Qy 118 PheArgAsnAspArg----- 122

Db 3531 GTGAATAACGATCGTAAAGCACTGTGATCCCTGACGCTGGCACTATCTAAATCCCCCG 3590

Qy 123 -----GlyArgPheLysAlaValAlaValCysValLeuSerAla 134

Db 3591 ACTTTGGTTTCCTTTGTTCCTCCCGACCAAACTTCAGTTGGCTCATGCTCTTGTGGCG 3650

Qy 135 Ile---ThrThrCysLeuAlaPheIleIleLysProAlaIleAsnAsnIleSerLeuMetIle 153

Db 3651 ATTGCTGGAGGGCTTGTGTGGTGAAGCCCATTTGTTACGGCTTTGTTCTCATGTTTC 3710

Qy 154 LeuGlyLeuProCysThrAlaLeuValAlaGluLeuLysArg 168

Db 3711 ATGAGTGTGGCCGACCATGTTAATGCTCTACACAGAGCTGCAGAGG 3755

RESULT 15

ABL30432

ID ABL30432 standard; DNA; 4071 BP.

XX ABL30432;

AC ABL30432;

XX

DT 26-MAR-2002 (first entry)

XX

DE *Drosophila melanogaster* genomic polynucleotide SEQ ID NO 42769.

XX *Drosophila*; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX *Drosophila melanogaster*.

OS

XX WO200171042-A2.

FN

XX 27-SEP-2001.

XX

XX 23-MAR-2001; 2001WO-US009231.

XX

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

XX (PEKE) PE CORP NY.

XX

XX Venter JC, Adams M, Li PWD, Myers EW;

XX

XX WPI; 2001-656860/75.

DR

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from *Drosophila* and for elucidating cell signaling and cell-cell interactions.

PT

PT

PT

PS Claim 1; SEQ ID NO 42769; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

CC

XX

SQ Sequence 4071 BP; 1027 A; 938 C; 941 G; 1165 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.55e-16	Length:	4071
Score:	264.50	Matches:	44
Percent Similarity:	70.83%	Conservative:	24
Best Local Similarity:	45.83%	Mismatches:	27
Query Match:	17.49%	Indels:	1
DB:	4	Gaps:	1

US-10-017-410-2 (1-275) x ABL30432 (1-4071)

Qy 168 ArgCysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeu 187

Db 3214 AGAGTTAGTACCAAGAGGCTACCGCTGGCGATCCGATCGACCGCTCTGGGCTGTT 3273

Qy 188 AlaLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPhe 207

Db 3274 GCGGTGTTCTGCTGGATCAATGACCGGATCTTTCGAGGCTGCTCTCGATCAATTTT 3333

Qy 208 ProTyrLeuHisCysValTrpHisIleLeuLeuCysLeuAlaSerTyrLeuGlyCysVal 227

Db 3334 CCATACCTGACGCGCTTCTGGCACATATTTCATTTTATAGCGGCTACACGGTGTGGTG 3393

Qy 228 CysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPhe 247

Db 3394 CTTTTCGCTACTTCTACGTGGAAATCGAGCTGCCGACGACGCGCTCTCGAAGTAC 3453

Qy 248 TrpProSerGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263

Db 3454 TGGCAAGAACGAGTTTCGAGTTC---GGGATACCCCTTCATTTCGATC 3498

Search completed: November 21, 2004, 07:51:29

Job time : 445 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 21, 2004, 07:04:58 ; Search time 98 Seconds
(without alignments)
1994.561 Million cell updates/sec

Title: US-10-017-410-2

Perfect score: 1512
Sequence: 1 MGAHPWHDLRAGSEVDWC.....IGVPYVSLCAHKKSPVKIT 275

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_D/US10017410/runat 19112004.133621.19022/app query.fasta.1.455
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410 @CGN 1.1.105 @runat 19112004.133621.19022 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1:	/cgn2_6/ptodata/1/ina/5A COMB.seq:*
2:	/cgn2_6/ptodata/1/ina/5B COMB.seq:*
3:	/cgn2_6/ptodata/1/ina/6A COMB.seq:*
4:	/cgn2_6/ptodata/1/ina/6B COMB.seq:*
5:	/cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6:	/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	209.5	13.9	1063	4	US-09-799-451-875
c 2	204	13.5	358	4	Sequence 875, App
3	173	11.4	915	4	Sequence 11241, A
c 4	118.5	7.8	6755	3	Sequence 1441, Ap
5	105.5	7.0	1257	4	Sequence 4, Appli
c 6	105.5	7.0	2859	4	Sequence 892, App
7	102	6.7	1575	4	Sequence 944, App
8	101.5	6.7	107820	4	Sequence 1139, Ap
9	101	6.7	1794	4	Sequence 1, Appli
c 10	98.5	6.5	3304	4	Sequence 577, App
11	97.5	6.4	960	4	Sequence 4182, Ap
12	97.5	6.4	92407	4	Sequence 2470, Ap
					Sequence 36, Appli

13	97	6.4	2304	4	US-09-540-236-426	Sequence 426, App
c 14	97	6.4	8155	4	US-08-956-171E-63	Sequence 63, Appl
c 15	97	6.4	8155	4	US-08-781-386A-63	Sequence 63, Appl
16	96.5	6.4	1297	3	US-09-083-521-4	Sequence 4, Appli
17	96.5	6.4	1863	1	US-09-198-284-4	Sequence 4, Appli
18	96.5	6.4	1863	2	US-08-987-122-4	Sequence 4, Appli
c 19	96.5	6.4	3010	4	US-08-961-527-25	Sequence 25, Appl
20	96	6.3	3267	3	US-09-453-702B-170	Sequence 170, App
21	95.5	6.3	5391	4	US-08-811-519-2	Sequence 2, Appli
22	95.5	6.3	5393	4	US-09-262-537-19	Sequence 19, Appl
23	95	6.3	1350	4	US-09-328-352-3355	Sequence 3355, Ap
24	94.5	6.2	3603	4	US-09-221-013A-5	Sequence 5, Appli
25	94.5	6.2	3673	4	US-09-221-013A-11	Sequence 11, Appl
26	93	6.2	774	4	US-09-252-991A-3254	Sequence 3254, Ap
27	93	6.2	5610	4	US-09-262-537-57	Sequence 57, Appl
28	93	6.2	246240	2	US-08-724-394A-20	Sequence 20, Appl
29	93	6.2	246240	2	US-08-724-394A-21	Sequence 21, Appl
30	93	6.2	246240	2	US-08-724-394A-22	Sequence 22, Appl
31	92.5	6.1	1329	4	US-09-602-787A-507	Sequence 507, App
32	92.5	6.1	1346	4	US-09-761-962A-12	Sequence 12, Appl
c 33	92.5	6.1	1806	4	US-09-248-796A-10630	Sequence 10630, A
34	92	6.1	648	4	US-09-583-110-774	Sequence 774, App
35	92	6.1	1571	4	US-09-221-017B-625	Sequence 625, App
c 36	92	6.1	1763	4	US-09-270-767-5952	Sequence 5952, Ap
c 37	92	6.1	1763	4	US-09-270-767-21234	Sequence 21234, A
c 38	92	6.1	51259	3	US-08-781-891-209	Sequence 209, App
c 39	92	6.1	51259	4	US-09-618-166-209	Sequence 209, App
40	91.5	6.1	1452	4	US-09-248-796A-6712	Sequence 6712, Ap
41	91.5	6.1	1733	4	US-09-560-761-17	Sequence 17, Appl
42	91.5	6.1	2612	4	US-09-270-767-13970	Sequence 13970, A
c 43	91.5	6.1	3897	4	US-09-540-236-414	Sequence 414, App
c 44	91.5	6.1	5099	4	US-09-887-052-1	Sequence 1, Appli
c 45	91.5	6.1	5099	4	US-09-887-052-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-799-451-875

; Sequence 875, Application US/09799451

; Patent No. 6783969

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan

; APPLICANT: Zhang, Jie

; APPLICANT: Xue, Aidong J.

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Ma, Yunqing

; APPLICANT: Yamazaki, Victoria

; APPLICANT: Chen, Rui-hong

; APPLICANT: Wang, Zhiwei

; APPLICANT: Wang, Dunrui

; APPLICANT: Yang, Yonghong

; APPLICANT: Wehrman, Tom

; APPLICANT: Ghosh, Reena

; APPLICANT: Drmanac, Radjoje T.

; TITLE OF INVENTION: No. 6783969el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 803

; CURRENT APPLICATION NUMBER: US/09/799,451

; CURRENT FILING DATE: 2001-03-05

; NUMBER OF SEQ ID NOS: 948

; SOFTWARE: pt_FL_genes Version 2.0

; SEQ ID NO 875

; LENGTH: 1063

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (2)..(802)

US-09-799-451-875

Alignment Scores:

Pred. No.: 1-5e-14 Length: 1063
Score: 209.50 Matches: 81
Percent Similarity: 40.89% Conservative: 29
Best Local Similarity: 30.11% Mismatches: 103
Query Match: 13.86% Indels: 58
DB: 4 Gaps: 7

US-10-017-410-2 (1-275) x US-09-799-451-875 (1-1063)

QY 14 SerSerGluValAspTrpCysGluAspAsnThrIleValProAlaIleAlaGluPhe 33
DB 44 ACCTCCACGCTGGACGTGGTGGAGGAACTACTCGTGACCTGGTACATCGCCGAGTTC 103
QY 34 TyrAsnThrIleSerAsnValLeuPheIleLeuProPheIleCysMetCysLeuPhe 53
DB 104 TGAATACAGTGAGTAAC---CTGATCATGATTATACCTCCAA--TGTTCCGTGCAATTC 158
QY 54 ArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTrpThrLeuValVal 73
DB 159 AGAGTGTAGAGACGCTCTGAAAGCGGTACATTGCTTCT-TATTTAGCACTCACAGTG 217
QY 74 ValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAsp 93
DB 218 GTAGGAATGGATCCCTGGTGTCTTCCATGACTCTGAAATATGAAATGACGCTATTGGAT 277
QY 94 GluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgGlyr 113
DB 278 GAATCCCAATGATATAC---AGCTGTTCATATTTGTGTACTGATGTTTGAATGTTTC 334
QY 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaValValCysValLeuSer 133
DB 335 AAGATCAAGAACTCAGTAACCTACCATCTGCTTTTACCTTAGTCTTCACTTCAATTAATA 394
QY 134 AlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsnAsnIleSerLeuMetIle 153
DB 395 GTAACACAGATTACCTTAAGGTAAAGAGCCCAATATCCATCAGTGTATGTAAGATG 454
QY 154 LeuGlyLeuProCysThrAlaLeuValAlaGluLeuLysArgCysAspAsnValArg 173
DB 455 TTG----- 457
QY 174 ValPheLysLeuGlyLeuPheSer----- 181
DB 458 GTCTTTACATTAGTACTTCGATCTAATTATATTGTATGTTTACATGGGTTTATCCATGGCTTAGA 517
QY 182 GlyLeuTrpTrpThr-----LeuAlaLeuPheCysTrpIleSerAsp 195
DB 518 GGACTGGGTATATACATCATTTGGGTATATTTTATTTGGGATTTTATTTTGGAAATATAGAT 577
QY 196 GlnAlaPheCysGluLeuSerSerPheHis----- 206
DB 578 AACATATTTGTGAGTCAGTCTGAGGAACTTTGGAAGAGTACCCTATCATAGGTATT 637
QY 207 PheProTyrLeuHisCysValTrpHisIleLeuIleCysLeuAlaSerTyrLeuGlyCys 226
DB 638 ACCACACAAATTTTCATGTCATGGTGGCATATTTTAACTGGCCTTGGTCTTCTATCTTCATC 697
QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArg 246
DB 698 CTTTTCAGTTTGTAT-----ACAAGAACACTTTTACCTGAGA 733
QY 247 PheTrpProSerGluLysTrpAlaPhe 255
DB 734 TATAGGCCAAAGTGAAGTTTCTCTTT 760

RESULT 2

US-09-621-976-11241/c

; Sequence 11241, Application US/09621976

; Patent No. 6639063

GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 11241
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-11241

Alignment Scores:
Pred. No.: 1-29e-14 Length: 358
Score: 204.00 Matches: 34
Percent Similarity: 97.30% Conservative: 2
Best Local Similarity: 91.89% Mismatches: 1
Query Match: 13.49% Indels: 0
DB: 4 Gaps: 0

US-10-017-410-2 (1-275) x US-09-621-976-11241 (1-358)

QY 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
DB 111 ATGGGGCGCCGACCTGGTGGGACCACTGCGAGCTGGTAGCTCGGAGGTGGACTGGTGC 52
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIle 37
DB 51 GAGGACAACTACCACTGCTGCTGCTATCGCCGAGTTCTTACAAACAGGTT 1

RESULT 3

US-09-248-796A-1441
; Sequence 1441, Application US/09248796A
; Patent No. 6747137

GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1441
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1441

Alignment Scores:
Pred. No.: 1-95e-10 Length: 915
Score: 173.00 Matches: 68
Percent Similarity: 40.64% Conservative: 34
Best Local Similarity: 27.09% Mismatches: 96
Query Match: 11.44% Indels: 54
DB: 4 Gaps: 12

US-10-017-410-2 (1-275) x US-09-248-796A-1441 (1-915)

QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
DB 85 ACTTCTACTACTGATTGGTGGAGAGAACTAGTCTATCTACATATATCGCAGAGCA 144
QY 34 TyrAsnThrIleSerAsnValLeuPheIleLeuProPheIleCys-----MetCys 51
DB 145 TTGAATACTACCACTAACTCAGTATTATTGCTTTTGGCAAC-ATTTGCGATATATCATGTC 203


```
US-10-017-410-2 (1-275) x US-09-792-616-1 (1-107820)
QY 19 TrpCysGluAspAsnThrThrileValProAlaIleAlaGluPheTyrAsnThrileSer 38
Db 38295 TTCTGTCCTTAAATGCTACTGTATG-----TTGCTGGAATTTGCCATCATGTACT 38348
QY 39 AsnValLeuPhePheIleLeuProPheIleCysMetCysLeuPheArgGlnTyrAlaThr 58
Db 38349 ATAATGTCGTGTTTCTTACTGTTACTATGCAATGCAATGCTATTCGTTACTATAATTC 38408
QY 59 CysPheAsnSerGlyIle-----TyrLeuIleTyrPheLeuValVal 73
Db 38409 AGTATTATTTGCTGGAATGTCATCATGTTACTATAATGCTGCTGTTTGTACTGTTA 38468
QY 74 ValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeu----- 88
Db 38469 -----CATTTGTTATGTTGCTACTCTGTTACTCTGTTACTGTTACAGT 38504
QY 89 -----GlyClnMetLeuAsp-----GluLeuAlaIleLeuTyrPheValLeuMetCys 103
Db 38505 CGTGTGCTGGGATGTTGCTGTCATGTTATGATGTTGCTGTTGTTGTTACTATTGCAATGC 38564
QY 104 AlaLeuAlaMetTrpPheProArgTyrLeuProLysIlePheArgAsnAspArgGly 123
Db 38565 TATGTTGCTGTTGTTGTTCTATCATGTTACTA-AAA----- 38599
QY 124 ArgPheLysAlaValValCysValLeuSerAlaIleThrThr-----CysLeu 139
Db 38600 -----TGTTGCTATTACGTTTACTATTATGCTATGTTGTTGTTGTTGTTATT 38644
QY 140 AlaPheIleLysProAla-----IleAsnAsnIleSerLeuMetIleLeuGlyLeuProCys 158
Db 38645 GCTGGAATGTCGCATCATGTTACTATAATGCTCTATTATGTTACTA---TTACACTGC 38701
QY 159 ---ThrAlaLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeu 177
Db 38702 TATGTTGCTATTCGTGA-----CTGCAATATGTTGTTATGTTGTTGTTGTTGTT 38746
QY 178 GlyLeuPheSerGlyLeuTyrTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAla 197
Db 38747 -----ACATTACTGTTTACTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTG 38782
QY 198 PheCysGluLeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeu 217
Db 38783 TTCTGGAATATTGCCAATGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCT 38836
QY 218 IleCysLeuAlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAaSerGlu 237
Db 38837 GCTATGTTG-----CTGTTGCAATGC---TGTTCTTGTGTTGCTGCTGCTGCTGCTG 38887
QY 238 Ile-ProGlu-----GlnGlyProValIleArgPheTrpProSerGluLysTr 253
Db 38888 GTGACCTGTTCTCCTCAAGCTCAGAACGAAAGAACCAAGATTGCGCTGCTCCAAA-- 38945
QY 253 palAPheIleGlyValProTyrValSerLeuLeuCysAlaHis 267
Db 38946 -----CCTATGCTCCTAACCACTGCTGCCAC 38972

RESULT 9
US-09-328-352-577
; Sequence 577, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ IDS NOS: 8252
; SEQ ID NO 577
; LENGTH: 1794

; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-577
Alignment Scores:
Pred. No.: 0.107 Length: 1794
Score: 101.00 Matches: 56
Percent Similarity: 37.97% Conservatives: 34
Best Local Similarity: 23.63% Mismatches: 67
Query Match: 6.68% Indels: 80
DB: 4 Gaps: 10

US-10-017-410-2 (1-275) x US-09-328-352-577 (1-1794)
QY 25 ThrileValProAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIle 44
Db 97 AGTATTATTTGCTGGAATGTCATCATGTTACTATAATGCTGCTGTTTGTACTGTTA 156
QY 45 LeuProPheIleCysMetCysLeuPheArgGlnTyrAlaThrCysPheAsnSer----- 62
Db 157 ATTCACACCTGTTATATTACTTATG-----ACAGCTGGGTGATGATCAAAA 201
QY 63 -----GlyIleTyrLeuIleTyrPheLeu 70
Db 202 AAGACTTTAAGTCAGGAGCTAAGGCAAACTTTACATGCGCATCTCGCGCTTAATGTCGGT 261
QY 71 LeuValValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGln 90
Db 262 TTGAGTTTGGCTGTTAATTTGGCAAGTTTATCAACCGGTGCAGACAGTATTATTTA--- 315
QY 91 MetLeuAspGluLeuAlaIleLeuTyrValLeuMetCysAlaLeuAlaMetTrpPhePro 110
Db 316 -----CTTTTCTGATTTGGACTTTGCTT----- 339
QY 111 ArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaValCys 130
Db 340 -----TTATGTCGGTGGCTATATCGCAAAATATTGGAATCTTT---GCTTTAATTTGT 390
QY 131 ValLeuSerAlaIleThrThrCysLeuAlaPheIleLysPro----- 144
Db 391 ATCACAAGCAACTTACACTTTTTTTTATTTTATGACAGACTTTTGGCGAGAGAAAGTTT 450
QY 145 -----AlaIleAsnAsnIleSerLeuMet-----IleLeuGly 155
Db 451 CCTATCTTTATCTTTTCTCTTAAATCTATTAAGCTTTGTGCGAGTTTGGATCTGTCAA 510
QY 156 LeuProCysThrAlaLeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPhe 175
Db 511 AAGAAATATACGGCTTTACGTTT-----ATTTT 540
QY 176 LysLeuGlyLeuPheSerGlyLeuTyrTrpThrLeuAlaLeuPheCysTrpIleSerAsp 195
Db 541 -----ATTGCTGTTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 567
QY 196 GlnAlaPheCysGluLeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHis 215
Db 568 ACAGGAATGATTCATTCATCAAGTGAATAATTTACCTTATCTG----- 612
QY 216 IleLeuIleCysLeuAlaSerTyrLeuGlyCysValCysPheAlaTyrPhe 232
Db 613 -----ATCTCGGCTTTTCTTAGGAATCATCGCTTTTATTATTATTTT 654

RESULT 10
US-09-710-279-4182/c
; Sequence 4182, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: F034800S
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 21, 2004, 07:44:08 ; Search time 467 Seconds
(without alignments)
3182.005 Million cell updates/sec

Title: US-10-017-410-2

Perfect score: 1512

Sequence: 1 MGAPWHDLRAGSEVDWC.....IGVPVSLCAHKKSPVKIT 275

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US10017410/runat_19112004_133621_19052/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10017410 @CIGN 1 1 723 @runat_19112004_133621_19052
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Match	Length	DB	ID	Description
1	1512	100.0	4175	13	US-10-017-410-1	Sequence 1, Appl
2	1433	94.8	4202	16	US-10-188-832-21	Sequence 21, Appl
3	1433	94.8	4212	15	US-10-295-027-151	Sequence 151, App
4	1404	92.9	828	13	US-10-017-410-3	Sequence 3, Appl
5	1373.5	90.8	822	15	US-10-182-447-6	Sequence 6, Appl
6	665	44.0	744	10	US-09-945-527-48	Sequence 48, Appl
7	549.5	36.3	792	15	US-10-182-447-4	Sequence 4, Appl
8	220.5	14.6	1194	16	US-10-425-114-32295	Sequence 32295, A
9	220.5	14.6	1471	18	US-10-425-115-107231	Sequence 107231, A
10	213.5	14.1	1547	17	US-10-767-701-12261	Sequence 12261, A
11	212.5	14.1	1436	17	US-10-437-963-42300	Sequence 42300, A
12	209.5	13.9	801	15	US-10-182-447-5	Sequence 5, Appl
13	209.5	13.9	1063	16	US-10-302-172-875	Sequence 875, App
14	206.5	13.7	1429	18	US-10-425-115-77897	Sequence 77897, A
15	186	12.3	35425	15	US-10-017-161-2429	Sequence 2429, Ap
16	186	12.3	35425	15	US-10-232-798-2069	Sequence 2069, Ap
17	185	12.2	1357	16	US-10-424-599-34549	Sequence 34549, A
18	165	10.9	275	18	US-10-425-115-111725	Sequence 111725, A
19	161	10.6	636	16	US-10-264-237-615	Sequence 615, App
20	127.5	8.4	2186	16	US-10-425-114-35305	Sequence 35305, A
21	127	8.4	3034	18	US-10-425-115-77895	Sequence 77895, A
22	123.5	8.2	1498	15	US-10-017-161-1999	Sequence 1999, Ap
23	123.5	8.2	1498	15	US-10-292-798-1645	Sequence 1645, Ap
24	119.5	7.9	568	17	US-10-767-701-30658	Sequence 30658, A
c	115	7.6	284	15	US-10-002-631C-17	Sequence 17, Appl
26	112.5	7.4	248436	13	US-10-087-192-2014	Sequence 2014, Ap
27	112	7.4	1501	15	US-10-017-161-2097	Sequence 2097, Ap
28	112	7.4	1501	15	US-10-232-798-1743	Sequence 1743, Ap
29	111	7.3	1880	14	US-10-198-846-12701	Sequence 12701, A
30	110.5	7.3	744802	15	US-10-292-798-1369	Sequence 1369, Ap
31	109	7.2	1282	15	US-10-017-161-2099	Sequence 2099, Ap
32	109	7.2	1282	15	US-10-292-798-1745	Sequence 1745, Ap
33	109	7.2	2485	14	US-10-198-846-13523	Sequence 13523, A
c	108.5	7.2	5842	11	US-09-984-429-316	Sequence 316, App
35	108	7.1	1984	15	US-10-017-161-2133	Sequence 2133, Ap
36	108	7.1	2614	13	US-10-232-798-1779	Sequence 1779, Ap
37	107.5	7.1	338702	13	US-10-087-192-292	Sequence 292, App
c	106.5	7.0	63828	13	US-10-087-192-388	Sequence 388, App
39	106	7.0	2195	18	US-10-425-115-92688	Sequence 92688, A
c	105.5	7.0	509	15	US-10-029-386-3107	Sequence 3107, Ap
41	105.5	7.0	439892	13	US-10-087-192-454	Sequence 454, App
42	105	6.9	1310	17	US-10-437-963-72944	Sequence 72944, A
c	104.5	6.9	53714	16	US-10-052-482-142	Sequence 142, App
c	104	6.9	2540	17	US-10-437-963-6583	Sequence 6583, Ap
c	102	6.7	876	13	US-10-001-857-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-10-017-410-1
; Sequence 1, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Graveel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; NUMBER OF SEQ ID NOS: 4
; CURRENT FILING DATE: 2001-12-14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4175
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)...(859)
US-10-017-410-1

Alignment Scores:

Pred. No.: 4,29e-157 Length: 4175
Score: 1512.00 Matches: 275
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-2 (1-275) x US-10-017-410-1 (1-4175)

QY 1 MetGlyAlaProHisTTPTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
DB 35 ATGGGCGCCCGCAGCTGGTGGACCACTCGGGCTGGCAGTTCGGAGGTGGATTGTGTC 94
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
DB 95 GAGGACAACTACACTATCGTGCCTGCCATTGCCGAGTTCACAAACGATCAGCAACGTC 154
QY 41 LeuPhePheIleLeuProProLleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
DB 155 TTGTTTTCATTTTACCTCCCATCTGCATGTGCTTGTTCGCCAGTAGTACGCAACGTGCTTC 214
QY 61 AsnSerGlyIleTyrLeuIleTyrThrLeuLeuValValGlyIleGlySerValTyr 80
DB 215 AACAGCGGCATCTACTTAATATGACCGCTCTAGTGTGTAGTGGGATGGATCTGTCTAC 274
QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
DB 275 TTCATATGCAACGGCTAGTTCCTGGGTGAGTGTGCTGATGCACTTGCATTCCTGTGGGT 334
QY 101 LeuMetCysAlaLeuAlaMetTyrPheProArgArgTyrLeuProLysIlePheArgAsn 120
DB 335 CTGATGTGTCTTGGCCATGTGGTTCCTCAGGAGGTATTACCAAGATCTTTCGGAAT 394
QY 121 AspArgGlyArgPheLeysAlaValValCysValLeuSerAlaIleTyrThrCysLeuAla 140
DB 395 GACAGGGCAGGTTCAGGCGAGTGGTGTGTCTGTCTGCTGCAATTACAGTGTGTGGCG 454
QY 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
DB 455 TTTATCAAGCCCGCATCAACATATTTCCCTGATGATTCTGGGACTTCCATGCACTGCG 514
QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
DB 515 CTGCTGTTCGAGAGCTGAAGAGGTGCAATGTGGTGTGTGTAAAGCTGGGCGCTCTTC 574
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGlu 200
DB 575 TCTGGCTCTGTGTGACTCTGGCTCTCTCTGCTGGATCAGCGACCAAGCCTTCTGTGAG 634
QY 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeuIleCysLeu 220
DB 635 CTGCTCTCTCTTCACTTCCCTACCTGCACTGTGTGTGGCATATTCCTCATCTGCTT 694
QY 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
DB 695 GCCTGTGACCTGGCTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 754
QY 241 GlnGlyProValIleArgPheTrpProSerGluLysTrpAlaPheIleGlyValProTyr 260
DB 755 CAAGGTCAGTCATCAGATTCGGCCCGCAGGAGAAATGGCTTTTATGTGTGCTCCCTAT 814
QY 261 ValSerLeuLeuCysAlaHisLysLysSerProValLysIleThr 275
DB 815 GTGTCCCTTCTGTGTGCCCCAAGAAGTCGCCAGTCAAGATCAG 859

RESULT 2

US-10-188-832-21

; Sequence 21, Application US/10188832

; Publication No. US20040076955A1

; GENERAL INFORMATION:

; APPLICANT: Mack, David H.

; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 21
; LENGTH: 4202
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-21

Alignment Scores:

Pred. No.: 2,63e-148 Length: 4202
Score: 1433.00 Matches: 256
Percent Similarity: 98.18% Conservatives: 14
Best Local Similarity: 93.09% Mismatches: 5
Query Match: 94.78% Indels: 0
DB: 16 Gaps: 0

US-10-017-410-2 (1-275) x US-10-188-832-21 (1-4202)

QY 1 MetGlyAlaProHisTTPTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
DB 77 ATGGGCGCCCGCAGCTGGTGGACCACTCGGGCTGGCAGTTCGGAGGTGGATTGTGTC 136
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
DB 137 GAGGACAACTACACCATCGTGCCTGCTATCGCGAGTTCACAAACGATCAGCAATGTC 196
QY 41 LeuPhePheIleLeuProProLleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
DB 197 TTAATTTTCAATTTACCGCCCATCTGCATGTGCTTGTTCCTCAGTAGTCAACATGCTTC 256
QY 61 AsnSerGlyIleTyrLeuIleTyrThrLeuLeuValValGlyIleGlySerValTyr 80
DB 257 AACAGTGGCATCTACTTAATCTGGACTCTTTTGTGTGTAGTGGAAATGGATCGTCTAC 316
QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
DB 317 TTCATATGCAACCTTAGTTCCTGGGTGAGTTCAGTTCAGTTCAGTTCCTTGGGTT 376
QY 101 LeuMetCysAlaLeuAlaMetTyrPheProArgArgTyrLeuProLysIlePheArgAsn 120
DB 377 CTGATGTGTCTTGGCCATGTGCTTCCCCAGAGGTATCTACCAAGATCTTTCGGAAT 436
QY 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleTyrThrCysLeuAla 140
DB 437 GACCGGGTAGGTTCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 496
QY 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
DB 497 TTTGTCAAGCCTGCCATCAACCAATCTCTCTGATGACCCCTGGGAGTTCCTTGCACGTCA 556
QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
DB 557 CTGCTCATCGCAGCTAAAGAGGTGTGACAACTGCTGTGTGTAAAGCTGGGCGCTCTTC 616
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGlu 200

Db 617 TCGGGCTCTGGTGGACCTGGCCCTGTCTGTGGATCAGTGACCGAGCTTCTGGGAG 676
Qy 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTyrPheHisLeuLeuLeuLeuLeu 220
Db 677 CTGCTGTATCTTCACTTCCCTTACCTGACCTGCTGATGGGACATCTCTGCTT 736
Qy 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluLeuProGlu 240
Db 737 GCTGCTACCTGGGCTGTGTATGCTTGGCTTACTTGTATGCTGCTCAGAGATCTCTGAG 796
Qy 241 GlnGlyProValIleArgPheTyrProSerGluLysTyrPheAlaPheIleGlyValProTyr 260
Db 797 CAAGGCTGTATCACTCAAGTCTGCGCCCAATGAGAAATGGGCTTCAATGGGTCCCTAT 856
Qy 261 ValSerLeuLeuCysAlaHisLysLysSerProValLysIleThr 275
Db 857 GTGTCCCTCTGTGTGGCAACAGAAATCATCAGTCAAGATCAG 901

RESULT 3

US-10-295-027-151
; Sequence 151, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 151
; LENGTH: 4212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-151

Alignment Scores:
Pred. No.: 2,64e-148 Length: 4212
Score: 1433.00 Matches: 256
Percent Similarity: 98.18% Conservative: 14
Best Local Similarity: 93.09% Mismatches: 5

Query Match: 94.78% Indels: 0
DB: 15 Gaps: 0
US-10-017-410-2 (1-275) x US-10-295-027-151 (1-4212)
Qy 1 MetGlyAlaProHisTyrTrpAspHisLeuArgHisLeuSerSerGluValAspTyrCys 20
Db 87 ATGGGCCCCCGGCACTGGTGGGACCACTGACAGCTGGTAGCTGGAGGTGGATGGTGGC 146
Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 147 GAGGACAACTACACCATCTGCTGCTATCGCCGAGTTCTACACACGATCAGCAATGTC 206
Qy 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
Db 207 TTATTTTTCATTTTACCGCCCATCTGCATGTGCTGTGTTGTCAGTAGTATCAACATGCTTC 266
Qy 61 AsnSerGlyIleTyrLeuIleTyrThrLeuLeuValValGlyIleGlySerValTyr 80
Db 267 AACAGTGGCATCTACTTAATCTGGACTCTTTTGGTTGTAGTGGAAATTGGAATCGTCTAC 326
Qy 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTyrVal 100
Db 327 TTCATGCAACCTTAGTTCTTGGGTAGATGCTGATGACTGAGCTCTTGGGTT 386
Qy 101 LeuMetCysAlaLeuAlaMetTyrPheProArgArgTyrLeuProLysIlePheArgAsn 120
Db 387 CTGATGTGTCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTGGGAAT 446
Qy 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThrCysLeuAla 140
Db 447 GACCGGGTAGGTTCAAGGTGGTGGTCTGCTGTGCGGTATCGAGTGGCTGGCA 506
Qy 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
Db 507 TTTGTCAAGCTGCCATCAACACATCTCTGTATGACCTGGGAGTCTCTTGACTGCA 566
Qy 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
Db 567 CTGCTCATCGCAGAGCTAAAGAGGTGTGACAACTGCGTGTGTTTAAAGCTGGGCTCTTC 626
Qy 181 SerGlyLeuTyrTrpThrLeuAlaLeuPheCysTyrPheSerAspGlnAlaPheCysGlu 200
Db 627 TCGGGCTCTGGTGGACCTGGCCCTGTCTGTGGATCAGTGCACCGAGCTTCTGGCAG 686
Qy 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTyrPheHisLeuLeuLeuLeu 220
Db 687 CTGCTGTATCTTCACTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
Qy 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db 747 GCTGCTACCTGGGCTGTGTATGCTTGGCTTACTTGTATGCTGCTGCTGCTGCTGCTGAG 806
Qy 241 GlnGlyProValIleArgPheTyrProSerGluLysTyrPheAlaPheIleGlyValProTyr 260
Db 807 CAAGGCTGTATCACTCAAGTCTGCGCCCAATGAGAAATGGGCTTCAATGGGTCCCTAT 866
Qy 261 ValSerLeuLeuCysAlaHisLysLysSerProValLysIleThr 275
Db 867 GTGTCCCTCTGTGTGGCAACAGAAATCATCAGTCAAGATCAG 911
RESULT 4
US-10-017-410-3
; Sequence 3, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Graveel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; CURRENT FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(825)
US-10-017-410-3

Alignment Scores:

Pred. No.: 4,15e-146 Length: 828
Score: 1404.00 Matches: 251
Percent Similarity: 96.73% Conservatives: 15
Best Local Similarity: 91.27% Mismatches: 9
Query Match: 92.86% Indels: 0
DB: 13 Gaps: 0

US-10-017-410-2 (1-275) x US-10-017-410-3 (1-828)

```
QY 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
Db 1 ATGGGGCGCCGACCTGGTGGAGCCAGCTGCAGGCTGGTAGCTCGGAGGTGGACTGGTGC 60
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 61 GAGGACAACTACACCACTGGCTGCTATCGCCGAGTTCTACAAACAGATCAGCAATGTC 120
QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
Db 121 TTATTTTTCATTTTACCGCCCATCTGATGCTGTGTTGATGAGTATGCAACATGCTTG 180
QY 61 AsnSerGlyIleTyrIleuIleTrpThrLeuValValGlyIleGlySerValTyr 80
Db 181 AACAGTGGACATCTACTTAATCTGACTCTTTGGTGTAGTGGGAATTGGATCCGCTAC 240
QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
Db 241 TTCCATTTTACCTTTAGTTTCTGGGTGAGTACCTTGATGAACTTGGAGTCCCTTGGGTT 300
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuProLysIlePheArgAsn 120
Db 301 CTGATGCTGTCTTGGCCATGTGTTCCCGAGAGGTATCTACCAAGATCTTTCCGAAT 360
QY 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThrCysLeuAla 140
Db 361 GACAGGGGTAGGTTCAGAGTGGTGGTGCAGTGTCTGTCTGCGGTACGAGTGCCTGGCA 420
QY 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
Db 421 TTGTGCAAGCTGCCATCAACACATCTCTGTGATGACCTGGGAGTTCCTTGCATGCA 480
QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
Db 481 CTGCTCATCGCAGAGCTAAGAGGTGTGACCAATCGGTGTGTTAAGCTGGGCCCTCTTC 540
QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTyrIleSerAspGlnAlaPheCysGlu 200
Db 541 TCGGGCCCTCTGGTGGACCTGGGCCCTGTCTGTGATGACCTGGGAGTTCCTTCTCGAG 600
QY 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeuIleCysLeu 220
Db 601 CTGTGTCATCTTCAACTTCCCTACCTGCACTGATGATGGGACATCTCATCTGCTT 660
QY 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db 661 GCTGCTACCTGGGCTGTGTATGCTTTGGCTACTTTGATGCTGCTCAGAGATTCCTGAG 720
QY 241 GlnGlyProValIleArgPheTrpProSerGluLysTrpAlaPheIleGlyValProTyr 260
Db 721 CAAGGCCCTGTCAATCAAGTTCTGGCCCAATGAGAAATGGGCCCTTCATGTGTGTCCTAT 780
QY 261 ValSerLeuLeuCysAlaHisLysLysSerProValIysIleThr 275
```

Db 781 GTGTCCCTCTGTGTGCCAACAGAAATCATCATGATCAAGATCAG 825

RESULT 5

US-10-182-447-6
; Sequence 6, Application US/10182447
; Publication No. US20030185814A1
; GENERAL INFORMATION:
; APPLICANT: HOFMANN, Kay
; APPLICANT: RADT, Marcus
; TITLE OF INVENTION: CERAMIDASE
; FILE REFERENCE: P68055US0
; CURRENT APPLICATION NUMBER: US/10/182,447
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: PCT/EP01/00900
; PRIOR FILING DATE: 2001-01-27
; PRIOR APPLICATION NUMBER: DE 10003293.1
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: DE 10011392.3
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-447-6

Alignment Scores:
Pred. No.: 1,01e-142 Length: 822
Score: 1373.50 Matches: 246
Percent Similarity: 94.91% Conservatives: 15
Best Local Similarity: 89.45% Mismatches: 13
Query Match: 90.84% Indels: 1
DB: 15 Gaps: 1

US-10-017-410-2 (1-275) x US-10-182-447-6 (1-822)

```
QY 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
Db 1 ATGGGGCGCCGACCTGGTGGAGCCAGCTGCAGGCTGGCAGCTCGGAGGTGGACTGGCGC 60
QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 61 GAGGACAACTACACCACTGGCTGCTGTGCTGCGCGAGTTCTATACATGATCAGCAATGTC 120
QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
Db 121 TTATTTTTCATTTTACCGCCCATCTGATGCTGTGTTGCTGATGATGCAACATGCTTC 180
QY 61 AsnSerGlyIleTyrIleuIleTrpThrLeuValValGlyIleGlySerValTyr 80
Db 181 AACAGGGGACATCTACTTAATCTGCG---CTCTTGGTGTAGCGGGAATTGGATCCGCTAC 237
QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
Db 238 TTCCATGCAACCTTAGTTTCTGGGTGAGTACCTTGATGAACTTGCAGTCCCTTGGGTT 297
QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuProLysIlePheArgAsn 120
Db 298 CTGATGCTGTGCTTCGGTTCATGTGTTCCCGAGAGGTATCTACCAAGATCTTTCCGAAT 357
QY 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThrCysLeuAla 140
Db 358 GACCAGGGTAGGTTCAGAGTGGTGGTGTGCTGCTGCTGCAAGTATGACGTCCTGGCA 417
QY 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
Db 418 TTGTGCAAGCTGCCATCAACACATCTCTGATGACCTCGGAGTTCCTTGGCTGCA 477
QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
Db 478 CTGCTCATCAGAGGCTAAAGAGGTGTGACCAATCGTGTGTTTAAAGCTGGGCCTCTTC 537
```

Qy 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGlu 200
Db TCGGGCTCTGGTGGACCTGGCCCTGTCTCTGGATCGATCAGTACCGAGCTTTCTGGGAG 597
Qy 201 LeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeuIleCysLeu 220
Db CTGCTGTCATCTCTCAACTTCCCTACTGTCACCTGTCATGTCATGTCATCTGCTT 657
Qy 221 AlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGlu 240
Db GCTGCTCTACTGGCTGTGTATGCTTGGCTACTTGTGCTGCTCAGAGATTCCTGAG 717
Qy 241 GlnGlyProValIleArgPheTrpProSerGluIleTrpAlaPheIleGlyValProTyr 260
Db CAAGGCCCTGTCAATCAATCTGGCCAGCGAGAAATGGGCCCTTCAITTTGGTGTCCCTAT 777
Qy 261 ValSerLeuLeuCysAlaHisLysLysSerProValIleThr 275
Db GTGTCTCTCTGTGTGCCAACAAGAAATCATCAGTCAAGACCAG 822
RESULT 6
US-09-945-527-48
; Sequence 48, Application US/09945527
; Publication No. US2003005588A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US2003005588A1el Nucleic Acid Molecules Encoding
; FILE REFERENCE: Nucleic Acid and Protein Homologs
; CURRENT APPLICATION NUMBER: US/09/945,527
; CURRENT FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-945-527-48

Alignment Scores:
Pred. No.: 5,096-64 Length: 744
Score: 665.00 Matches: 122
Percent Similarity: 94.74% Conservative: 4
Best Local Similarity: 91.73% Mismatches: 5
Query Match: 43.98% Indels: 2
DB: 10 Gaps: 1

US-10-017-410-2 (1-275) x US-09-945-527-48 (1-744)
Qy 1 MetGlyAlaProHisTrpTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
Db 277 ATGGGCGCCGCACTGGTGGACAGCTGCAGCTGTGTAGCTCGAGGTGAGTGTGTC 336
Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
Db 337 GAGGACAACTACACCATCGTCTGCTATCGCCGAGTTCTACACACGATCAGCAATGTC 396
Qy 41 LeuPhePheIleLeuProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
Db 397 TTATTTTCATTTTACCGCCCATCTGATGTGCTTTGTGTAGTGGAAATGGATCGCTAC 456
Qy 61 AsnSerGlyIleTyrLeuIleTrpThrLeuLeuValValGlyIleGlySerValTyr 80
Db 457 AACAGTGGCATCTACTTAATCTGGACTCTTTTGTGTGTAGTGGAAATGGATCGCTAC 516
Qy 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
Db 517 TTCCATGCAACCCCTAGTTTCTGGGTGAGATGCTTGTATGAATTCGAGTCCCTTTGGGT 576
Qy 101 LeuMetCysAlaLeuAlaMetTrpPheProArgGlyTyrLeuProIlePheArgAsn 120
Db 577 CTGATGTGTCTTGGCCATGTGTTCCCAAGAGGTATCTACCAAGATCTTTCCGAAT 636

Qy 121 AspArgGlyArgPheLysAlaValValCysValLeuSer 133
Db 637 GAC-----CGATACCCACCACTCAGTCACCATGATGCTTCT 669
RESULT 7
US-10-182-447-4
; Sequence 4, Application US/10182447
; Publication No. US2003018581A1
; GENERAL INFORMATION:
; APPLICANT: HOFMANN, Kay
; APPLICANT: RADT, Marcus
; TITLE OF INVENTION: CERAMIDASE
; FILE REFERENCE: P68055USO
; CURRENT APPLICATION NUMBER: US/10/182,447
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: PCT/EP01/00900
; PRIOR FILING DATE: 2001-01-27
; PRIOR APPLICATION NUMBER: DE 10003293.1
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: DE 10011392.3
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-447-4
Alignment Scores:
Pred. No.: 3,86e-51 Length: 792
Score: 549.50 Matches: 103
Percent Similarity: 59.38% Conservative: 49
Best Local Similarity: 40.23% Mismatches: 102
Query Match: 36.34% Indels: 2
DB: 15 Gaps: 1

US-10-017-410-2 (1-275) x US-10-182-447-4 (1-792)

Qy 9 HisLeuArgAlaGly-SerSerGluValAspTrpCysGluAspAsnTyrThrIleValPr 28
Db 9 CATCTTCGCTATCAGAGCTCCGAGGTGGACTGGTGTGAGAGCACTTCAGTACTCGGA 68
Qy 28 oAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPheIleLeuProIle 48
Db 69 GCTGTGCGCGAGTTCTACACAGCTTCTCAATATCCCTTCTTCTTCTTCTTCTTCTTCT 128
Qy 48 eCysMetCysLeuPheArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTr 68
Db 129 GATGATGCTCTGATGACCCGCTATGCCAGAGCGCTCCGCTACATTTACGTTGTCTG 188
Qy 68 pThrLeuLeuValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLe 88
Db 189 GTCCTCTTCATGATCATGAGCTTCTTCATGATTTCCATGATGAGCTCAGCTTCTCT 248
Qy 88 uGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTr 108
Db 249 GGGCCAGCTGCTGGACGAGATCGCATCTCTGCTCTCTGGCAGTGGCTATAGCATATG 308
Qy 108 pPheProArgGlyTyrLeuProIlePheArgAsnAspArgGlyArgPheLysAlaVal 128
Db 309 GATGCCCGCTGCTATTTCCCTCTCTCTTGGGGGAGACAGGTCCAGTTCATCCGCT 368
Qy 128 lValCysValLeuSerAlaIleThrCysLeuAlaPheIleIlePheProAlaIleAsnAs 148
Db 369 GGTCTTTCATCACCCTGCTGTCAGACCTTCTGCTCTCTGCGCCACAGCTCAAGC 428
Qy 148 nIleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeuLysAr 168
Db 429 CTAGCCCTCAACAGCATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 488
Qy 168 gCysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAl 188

```
Db 489 GACCACATAAGGACCTTCGGACACCTGATTGAGGTCTCGTGGTTTATGGGCTTTGC 548
QY 188 aleuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPhePr 208
Db 549 TCTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 608
QY 208 oTyrLeuHisCysValTrpHisIleLeuIleCysLeuAlaSerTyrLeuGlyCysValCy 228
Db 609 CTAATCTGCACAGCATCTGGCAATGCTCATCAGCATCACCTCCCTTATGGCATGTAC 668
QY 228 sPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPheTr 248
Db 669 CATGGCTTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
QY 248 pProSerGluLysTrpAlaPheIleGlyValProTyrValSerLeu 263
Db 729 GCCTCGGACAGTTGGCCC---GTGGGGCTGCCTACGTGGAATC 771

RESULT 8
US-10-425-114-32295
; Sequence 32295, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32295
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73286S07_FLI
US-10-425-114-32295

Alignment Scores:
Pred. No.: 2,64e-14 Length: 1194
Score: 220.50 Matches: 76
Percent Similarity: 43.73% Conservatives: 39
Best Local Similarity: 28.90% Mismatches: 121
Query Match: 14.58% Indels: 27
DB: 16 Gaps: 9

US-10-017-410-2 (1-275) x US-10-425-114-32295 (1-1194)
QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 117 ACATCAACGACTGAGTTGGAGGAGATTTATGACACTCGTCATATATCGAGAAATTC 176
QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuLeuProIleCysMetCysLeuPhe 53
Db 177 TACAATACCATCTCTAATGTCCCATGCTTTTGGCCTTATT-----GGATTAGTG 230
QY 54 ArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTrpThrLeuLeuValVal 73
Db 231 AATGCTTCTGCAAGTTTGGAGAACATTTAGTGTCTGACATATCCAAATATGATA 290
QY 74 ValGlyIleGlySerValTyrPheHisAlaThrLeuSerPhePheGlnMetLeuAsp 93
Db 291 CTTCGCTATTGGGAGTATGATTTTCCATGCGCACCTTCCCACTTCTCCTACACAGAGTGAT 350
QY 94 GluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMet-----Trp 108
Db 351 GAGACTCCTATGTTGGGAGATCTCTCTCATCATGATGATGCTCTCTATTACCGGACTGG 410
```

```
QY 109 PheProArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaVal 128
Db 411 CACTACAGGAGACAAATGCCA---ACTTTCCTGTTCTCTGTATGCTGCTGCTGTCGCGTA 467
QY 129 ValCysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsnAsn 148
Db 468 GTACATTTCTTTCGCCGGTTCCAAAGTCGTATTC-----AAGCTGCATTAC 512
QY 149 IleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeuLysArg 168
Db 513 GTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 572
QY 169 CysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAla 188
Db 573 GTGGGTGCGAAGCGGCTCGCAAACTGTGGTC-----CTTACATGACCTTGGG 623
QY 189 LeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPhe--- 207
Db 624 ACCCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 683
QY 208 ProTyrLeuHisCysValTrpHisIleLeuIleCysLeuAlaSerTyrLeuGlyCysVal 227
Db 684 CGCAGGGGCGACGCGTGGTGGCAGCTGCTTATGGGCTCAACTCGTACTATGCAAAACACG 743
QY 228 CysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPhe 247
Db 744 ---TTCTGATGTTCTGCGGGCTCAGCAG-----CCGGGG 776
QY 248 TrpProSerGluLysTrpAlaPheIleGly---ValProTyrValSerLeuLeuCysAla 266
Db 777 TGGGAGCGCGGATCAGCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836
QY 267 HisLysLys 269
Db 837 GAGAAGAGG 845

RESULT 9
US-10-425-115-107231
; Sequence 107231, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 107231
; LENGTH: 1471
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_29295C.1
US-10-425-115-107231

Alignment Scores:
Pred. No.: 3.57e-14 Length: 1471
Score: 220.50 Matches: 76
Percent Similarity: 43.73% Conservatives: 39
Best Local Similarity: 28.90% Mismatches: 121
Query Match: 14.58% Indels: 27
DB: 18 Gaps: 9

US-10-017-410-2 (1-275) x US-10-425-115-107231 (1-1471)
QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 303 ACATCAACGACTGAGTTGGAGGAGATTTATGACACTCGTCATATATCGAGAAATTC 362
```

QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
Db 363 TACAATACCATCTCTAAATGTCCCAATGCGCTCTTTGGCACTTATT-----GGATTAGT 416
QY 54 ArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThrLeuLeuValVal 73
Db 417 AATGCTTCTGCCAAGGTTTGAGAAAGATTTAGTGTCTGACATATCCAAATATGATA 476
QY 74 ValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAsp 93
Db 477 CTGTGATTATGGAGTATGATTTTCCATGCCACCTTTGCAACTTCTCTCAACAGAGTAT 536
QY 94 GluLeuAlaIleLeuTyrValLeuMetCysAlaLeuAlaMet-----Tyr 108
Db 537 GAGACTCTCTGTTGGAGATCTCTCTACATGTATGCTCTATTATCCCGACTGG 596
QY 109 PheProArgGlyTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaVal 128
Db 597 CACTACAGAGCACAATGCCA---ACTTTCCTGTTCCTGTATGGTGTGCTTGGCGTA 653
QY 129 ValCysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsnAsn 148
Db 654 GTACATTTCTTGGCCGGTTCACAGTGTATTTC-----AAGCTGCATTAC 698
QY 149 IleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGlnLeuLysArg 168
Db 699 GTTGGCTCTCTGCTCTGATCCCGCGATGTACAAAGTACTACATACAGCAAGAAC 758
QY 169 CysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTyrThrLeuAla 188
Db 759 GTGGTGGCAAGCGCTCGCAAACTGTGGGTC-----CTTACACTGACCTTGGG 809
QY 189 LeuPheCysTrpIleSerAspGlnAlaPheCysGlyLeuLeuSerSerPheHisPhe--- 207
Db 810 ACCCTCTGCTGCTGTTGTATGCTGCTCTTGCAGAAAGCTTTGGCAATTGGTACGTCAAC 869
QY 208 ProTyrLeuHisCysValTyrPheIleLeuLeuCysLeuAlaSerTyrLeuGlyCysVal 227
Db 870 CCGCAGGGCGCAGCGGTGGTGACGCTGTATTATGGCCCTCAACTGTACTATGCAAAACACG 929
QY 228 CysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPhe 247
Db 930 ---TTCCTGATGTTCTGCGGCTCAGCAG-----CGCGGG 962
QY 248 TrpProSerGluLysTrpAlaPheIleGly---ValProTyrValSerLeuLeuCysAla 266
Db 963 TGGGAGCGCGGATCAGCAGCCTCTCTGGATTCTTGGCGGTATGTCAGAGTCCAGAAACCA 1022
QY 267 HisLysLys 269
Db 1023 GAGAAGAGG 1031

RESULT 10
US-10-767-701-12261
; Sequence 12261, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 12261
; LENGTH: 1547
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS34423_1

US-10-767-701-12261

Alignment Scores:
Pred. No.: 2,31e-13 Length: 1547
Score: 213.50 Matches: 79
Percent Similarity: 43.56% Conservative: 36
Best Local Similarity: 29.92% Mismatches: 120
Query Match: 14.12% Indels: 29
DB: 17 Gaps: 10

US-10-017-410-2 (1-275) x US-10-767-701-12261 (1-1547)

QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 613 ACGTCAACAGCTAGATTGTGGAGAGAAATTATGCACACTCATCATATATCGCAATTC 672
QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCys---Leu 52
Db 673 TACAATACCATCTCTAATGTCCATGCGTCTCTTTGGCGGTATTGGATTAGTGAATGCT 732
QY 53 PheArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThrLeuLeuVal 72
Db 733 TTCGCCAA-----GGTTTGGAGAAACGGTTCAGTGTCTTCACGTATCCAATATG 783
QY 73 ValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeu 92
Db 784 ATACTGTCTATTGGAGTATGATTTTCCATGCCACCTTGCACACTTCTCTCCACAGAGT 843
QY 93 AspGluLeuAlaIleLeuTyrValLeuMetCysAlaLeuAlaMet----- 107
Db 844 GATGAGACTCCGATGGTCTGGAGATCTCTTATATATGTATGCTTTATTTCACCGGAC 903
QY 108 TrpPheProArgGlyTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysAla 127
Db 904 TGGCACTACAGAGCACAATGCCA---ACTTTCCTTTCTCTATATGGTGTGCTGCTTTGCA 960
QY 128 ValValCysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsn 147
Db 961 GTAGTTCAATTTCTTGGCCGGTTCACAGTTGTATTTC-----AAGTTGCAT 1005
QY 148 AsnIleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeuLys 167
Db 1006 TACATTGTCTGTGCTCTCTGTCATCCCGGAGTACAAAGTACTACATACGAGAA 1065
QY 168 ArgCysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTyrTrpThrLeu 187
Db 1066 GACATGGCGGCGAAGCGCTTGCAGAAAGCTTTGGGT-----CTTACATTAACCTT 1116
QY 188 AlaLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSerSerPheHisPhe 207
Db 1117 GGGACTCTCTGCTGCTAGTTGATGCGCATCTCTGTAAAGAGCTTTCATATGTTATGTC 1176
QY 208 ---ProTyrLeuHisCysValTyrPheIleLeuIleCysLeuAlaSerTyrLeuGlyCys 226
Db 1177 AACCGCAGCGGCGACGATGCTGGCATGTCTTATGGCCCTCAACTCATATATGCAAA 1236
QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArg 246
Db 1237 ACA---TTCCTAATGTTTGTGGCGGCTCAGCAA-----CGT 1269
QY 247 PheTrpProSerGluLysTrpAlaPheIleGly---ValProTyrValSerLeuLeuCys 265
Db 1270 GGGTGGGAGCGCGGATCACAACCTTCTTGGATTCTTGGCTTACGTCAGGTCAGAA 1329
QY 266 AlaHisLysLys 269
Db 1330 CCAGAAAGAGG 1341

RESULT 11

US-10-437-963-42300
; Sequence 42300, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 42300
LENGTH: 1436
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_45566C.1
US-10-437-963-42300

Alignment Scores:
Pred. No.: 2,68e-13 Length: 1436
Score: 212.50 Matches: 73
Percent Similarity: 42.70% Conservative: 41
Best Local Similarity: 27.34% Mismatches: 118
Query Match: 14.05% Indels: 35
DB: 17 Gaps: 10

US-10-017-410-2 (1-275) x US-10-437-963-42300 (1-1436)

Qy 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 260 ACTTCAACACTGAGCTCTCGAGGAAATATGACACTCGTCATATATGCGAGATTC 319
Qy 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProPheIleCysMetCysLeuPhe 53
Db 320 TACAACTACTGTTCTAATGTT-----CGATGTGTTCTCTGGCACTT 361
Qy 54 ArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTrpThrLeuLeu----- 71
Db 362 GTTGGACTTGTGAATGCTCTTCGCAAGGTTTGTGAGAAACGATTTCAGTGTCTGCACATA 421
Qy 72 -----ValValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGly 89
Db 422 TCCAAATATGATCTTGTCTCGGAGTATGATCTTCCATGCCACTTACAGCAGTTTGA 481
Qy 90 GlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMet----- 107
Db 482 CAGCAGAGCATGAGACTCCCAATGGTGTGGAGATTTCTCTATATCTTTATGTACTTTAT 541
Qy 108 -----TrpPheProArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArg 124
Db 542 TCACCAGACTGGCATTACCGGACGACATGCTACTTCTCTT-----TTCCCTA 589
Qy 125 PheLysAlaValValCysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysPro 144
Db 590 TAGGTCCTGCTTTTGCAGTAGTCACTTCTCTGTCGATCCAGGATGGGTATTCAG--- 646
Qy 145 AlaIleAsnIleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAla 164
Db 647 ---TTGCATTACGTTGGCTCTGCTCTCTATGCATCCAGGATGTACAGTACTACATA 703
Qy 165 GluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTrp 184
Db 704 CAAACTAAAGACATGGCTGCCAAGCGCTAGCAAGCTGTGGTT-----CTTACG 754
Qy 185 TrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSer 204
Db 755 TTGAGCTTGGCGACTCTTTGCTGGCTATTGATCGAATGTTCTGTAAGAGCTTTTCAT 814
Qy 205 PheHisPhe---ProTyrLeuHisCysValTrpHisIleLeuLeuIleCysLeuAlaSerTyr 223

Db 815 TGGTACGTCAACCCACAGGGCAGCATGTTGGCATATTCTCATGGGATTTAACTCATAC 874
Qy 224 LeuGlyCysValCysPheAlaTyrPheAspAlaIleSerGluIleProGluGlnGlyPro 243
Db 875 TTTGCAACACA---TTCTTAATGTTTTCGCGAGCTCAACAG----- 913
Qy 244 ValIleArgPheTrp---ProSerGluLysTrpAlaPheIleGlyValProTyrValSer 262
Db 914 -----CGTGGGTGGAGCCCAAAATACCACCTTTTCGGGTCTTTCGCTTATGTCAAG 967
Qy 263 LeuLeuCysAlaHisLysLys 269
Db 968 ATTCAAGAACCCCAAAAGAGG 988

RESULT 12
US-10-182-447-5
Sequence 5, Application US/10182447
Publication No. US20030185814A1
GENERAL INFORMATION:
APPLICANT: HOFMANN, Kay
APPLICANT: RADT, Marcus
TITLE OF INVENTION: CERAMIDASE
FILE REFERENCE: P68055USO
CURRENT APPLICATION NUMBER: US/10/182,447
CURRENT FILING DATE: 2002-07-29
PRIOR APPLICATION NUMBER: PCT/EP01/00900
PRIOR FILING DATE: 2001-01-27
PRIOR APPLICATION NUMBER: DE 10003293.1
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: DE 10011392.3
PRIOR FILING DATE: 2000-03-09
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 801
TYPE: DNA
ORGANISM: Homo sapiens
US-10-182-447-5

Alignment Scores:
Pred. No.: 2,47e-13 Length: 801
Score: 209.50 Matches: 81
Percent Similarity: 40.89% Conservative: 29
Best Local Similarity: 30.11% Mismatches: 103
Query Match: 13.86% Indels: 58
DB: 15 Gaps: 7

US-10-017-410-2 (1-275) x US-10-182-447-5 (1-801)

Qy 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
Db 43 ACCTCCAGCTGGAGTGTGCGAGGAACTACTCCGTGACCTGGTGCATCGCGCGAGTTC 102
Qy 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProPheIleCysMetCysLeuPhe 53
Db 103 TGGAAATACATGAGTAAC---CTGATCATGATATATACCTCAA---TGTTCGGTGCATTC 157
Qy 54 ArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTrpThrLeuValVal 73
Db 158 AGAGTGTGTAGACGGTCTGGAAGGCGGTACATTTGCTCT-TATTTAGCACTCACAGTG 216
Qy 74 ValGlyIleGlySerValTyrPheHisAlaThrIleSerPheLeuGlyGlnMetLeuAsp 93
Db 217 GTAGGAATGGATCCTCGTCTCCACATGACTCTGAAATATGAAATGCACTATTGGAT 276
Qy 94 GluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyr 113
Db 277 GAATCCCAATGATATAC---AGCTGTTCATATTTGTGTACTGTCATGTTTGAATGTTTC 333
Qy 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaValValCysValLeuSer 133
Db 334 AAGATCAAGAACTCAGTAAACTACCACTCTGCTTTTACCTTAGTCTTATTACAGTTTAATA 393

QY 134 AlaIleThrThrCysLeuAlaPheIleLeuSerProAlaIleAsnAsnIleSerLeuMetIle 153
 Db 394 GTAACACAGTTTACCTTAAAGTAAAGAGCCCAATATTCATCAGGTGATGGAATG 453
 QY 154 LeuGlyLeuProCysThrAlaLeuValAlaGluLeuLysArgCysAspAsnValArg 173
 Db 454 TTG----- 456
 QY 174 ValPheLeuGlyLeuPheSer----- 181
 Db 457 GTCTTTACATTAGTACTTCGACTATTATATTATTATGTTTACATGGGTTTATCCATGGCTTAGA 516
 QY 182 GlyLeuTyrTrpThr-----LeuAlaLeuPheCysTrpIleSerAsp 195
 Db 517 GGACTGGGTTATACATCATCTGGGTATATTTTATTGGGATTTTATTGGGAATATAGAT 576
 QY 196 GlnAlaPheCysGluLeuLeuSerSerPheHis----- 206
 Db 577 AACATATTGTGAGTCACTGAGGAACTTTCGAAGAAGGTACCACCTATCATAGGTATT 636
 QY 207 PheProTyrLeuHisCysValTrpHisIleLeuLeuLeuCysLeuAlaSerTyrLeuGlyCys 226
 Db 637 ACCACACAATTTTCATGTCATGGTGGCATATTTTAACTGGGCTTGGTTCCTATCTTCACATC 696
 QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArg 246
 Db 697 CTTTTCAGTTTGTAT-----ACAAGAACACTTTTACCTGAGA 732
 QY 247 PheTrpProSerGluLysTrpAlaPhe 255
 Db 733 TATAGGCCCAAAAGTGAAGTTTCTCTTT 759

RESULT 13

US-10-302-172-875
 ; Sequence 875, Application US/10302172
 ; Publication No. US20040053250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 803.1CNCB
 ; CURRENT APPLICATION NUMBER: US/10/302,172
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/225,251
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: PCT US02/05095
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 09/799,451
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 950
 ; SOFTWARE: pt FL Genes Version 2.0
 ; SEQ ID NO 875
 ; LENGTH: 1063
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (2)..(802)
 US-10-302-172-875

Alignment Scores:

Pred. No.:	3.72e-13	Length:	1063
Score:	209.50	Matches:	81
Percent Similarity:	40.89%	Conservative:	29
Best Local Similarity:	30.11%	Mismatches:	103
Query Match:	13.86%	Indels:	58
DB:	16	Gaps:	7

US-10-017-410-2 (1-275) x US-10-302-172-875 (1-1063)

QY 14 SerSerGluValAspTrpCysGluAspAsnTyrThrIleValProAlaIleAlaGluPhe 33
 Db 44 ACCTCCAGCTGGTGGTGGAGGAACTACTCGTACCTGGTACATCGCGGATTC 103
 QY 34 TyrAsnThrIleSerAsnValLeuPhePheIleLeuProProIleCysMetCysLeuPhe 53
 Db 104 TGAATACAGTGGTAAC---CTGATCATGATTATACCTCCAA--TGTTCGGTGGCAATTC 158
 QY 54 ArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThrLeuLeuValVal 73
 Db 159 AGAGTGTGTAGACGGTCTGGAAGAGCGGTACATTCCTTCT-TATTACACTCAGTG 217
 QY 74 ValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAsp 93
 Db 218 GTAGGAATGGGATCCTGGTCTTCCACATGACTCTGAAATATGAAATGAGCTATTGGAT 277
 QY 94 GluLeuAlaIleLeuTyrValLeuMetCysAlaLeuAlaMetTyrPheProArgTyr 113
 Db 278 GAATCCCATGATATAC---AGCTGTGCTATTTGTGTACTGCGATGTTTGAATGTTTC 334
 QY 114 LeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaValValCysValLeuSer 133
 Db 335 AAGATCAAGAAGTCACTAGTAACTACCATCTGCTTTTACCTTAGTCTTATTCAGTTAATA 394
 QY 134 AlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsnAsnIleSerLeuMetIle 153
 Db 395 GTAACACAGTTTACCTTAAAGGTAAAGAGCCCAATATTCATCAGTCTATGATGGAATG 454
 QY 154 LeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeuLysArgCysAspAsnValArg 173
 Db 455 TTG----- 457
 QY 174 ValPheLysLeuGlyLeuPheSer----- 181
 Db 458 GTCTTTACATTAGTACTTCGATCTATTATTATTATTGTTTACATGGGTTTATCCATGGCTTAGA 517
 QY 182 GlyLeuTyrTrpThr-----LeuAlaLeuPheCysTrpIleSerAsp 195
 Db 518 GGACTGGGTTATACATCATTTGGGTATATTTTATTGGGATTTTATTGGGAATATAGAT 577
 QY 196 GlnAlaPheCysGluLeuLeuSerSerPheHis----- 206
 Db 578 AACATATTGTGAGTCACTGAGGAACTTTCGAAGAAGGTACCACCTATCATAGGTATT 637
 QY 207 PheProTyrLeuHisCysValTrpHisIleLeuLeuCysLeuAlaSerTyrLeuGlyCys 226
 Db 638 ACCACACAATTTTCATGTCATGGTGGCATATTTTAACTGGGCTTGGTTCCTATCTTCACATC 697
 QY 227 ValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArg 246
 Db 698 CTTTTCAGTTTGTAT-----ACAAGAACACTTTTACCTGAGA 733
 QY 247 PheTrpProSerGluLysTrpAlaPhe 255
 Db 734 TATAGGCCCAAAAGTGAAGTTTCTCTTT 760

RESULT 14

US-10-425-115-77897
 ; Sequence 77897, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 77897
 ; LENGTH: 1429

RESULT 15

```

Db      21352  ||| .....:
Qy      75  -----GlyIleGlySe 78
Db      21412  GGGTTTAGGAGGTGGCGGACCCACCTGACCGCCTGCCCTTGCCGCTGCAGGCCTGTCTC 21471
Qy      78  rValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLe 98
Db      21472  CATGTATTTCCACATGACGCTCAGCTTCCTGGGCCAGCTGCTGGACGAGATGCCATCCT 21531
Qy      98  uTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePh 118
Db      21532  GTGGCTCCTGGCGAGTGGCTATAGCATATGGATGCCCGCTGCTATTCCCTCCTTCTCT 21591
Qy      118  eArgAsnAspArg 122
Db      21592  TGGGGGGAACAGG 21604

```

Search completed: November 21, 2004, 09:54:07
 Job time : 482 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 21, 2004, 06:57:39 ; Search time 2986 Seconds
(without alignments)
3355.970 Million cell updates/sec

Title: US-10-017-410-2

Perfect score: 1512

Sequence: 1 MGAPHWDLRAGSEVDWC.....IGVPVSLCAHKKSPVKIT 275

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10017410/runat_19112004_133620_19010/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -MODE=LOCAL
-DOCALIGN=200 -THR_SCORE=spt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=spt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10017410@cgn 1 1 5180 @runat_19112004_133620_19010 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1392	92.1	1022	5	BU511164 AGENCOURT
2	1230	81.3	681	9	AY401891 Mus muscu
3	1176	77.8	868	6	CA976684 AGENCOURT
4	1176	77.8	1173	3	AK085306 Mus muscu
5	1163	76.9	681	9	AY401890 Pan trogl
6	1158.5	76.6	797	6	CA463294 AGENCOURT
7	1157	76.5	681	9	AY401889 Homo sapi
8	1110.5	73.4	1527	3	AF370405 Homo sapi
9	1001	66.2	698	5	BU234223 603792086

10	996	65.9	565	6	CF115220	Shultzomi
11	986	65.2	633	2	BB660847	BB660847
12	984	65.1	622	6	CB723138	UI-M-GHO-
13	964	63.8	781	5	BU232554	603409105
14	920	60.8	548	6	CF169369	CF169369 B0812G07-
15	908	60.1	648	4	RJ060795	RJ060795
16	904	59.8	937	5	BX842808	BX842808
17	896	59.3	847	5	BU220481	603107516
18	855	56.5	678	5	BU203269	BU203269 604153603
19	834	55.2	498	4	BI848265	BI848265 470659 MA
20	773	51.1	432	6	CF169808	B0818D08-
21	766	50.7	449	5	BX646596	BX646596 DKF2781B
22	715	47.3	454	6	CB784796	AMGNNUC.T
23	692	45.8	401	2	BF554219	UI-R-CO-H
24	691.5	45.7	666	5	BW215750	BW215750
25	689.5	45.6	725	5	BW442024	BW442024
26	684.5	45.3	803	7	CF593785	AGENCOURT
27	669	44.2	427	1	AA900336	UI-R-E0-C
28	663	43.8	434	6	CB758986	AMGNNUC.M
29	653.5	43.2	815	5	BX623644	BX623644
30	651	43.1	399	6	CB706492	AMGNNUC.M
31	644.5	42.6	741	5	BW216493	BW216493
32	629.5	41.6	480	6	CB728612	AMGNNUC.M
33	628.5	41.6	696	5	BW351328	BW351328
34	625	41.3	485	2	BF549345	UI-R-A0-a
35	592	39.2	423	6	CA967276	CA967276
36	587	38.8	646	6	CA969385	CCLX06a22
37	575.5	38.1	712	5	BX620332	BX620332
38	570.5	37.7	398	6	CB707581	AMGNNUC.M
39	566	37.4	847	7	CN060612	CN060612
40	563	37.2	771	7	CN057662	Salamande
41	562	37.2	634	1	AL864302	AL864302
42	558	36.9	723	4	BM602328	BM602328 170006870
43	552	36.5	357	5	BY168309	BY168309
44	547	36.2	685	4	BJ062108	BJ062108
45	541	35.8	365	5	BY117228	BY117228

ALIGNMENTS

RESULT 1
LOCUS BU511164 1022 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT_10107530 NIH MGC 134 Mus musculus cDNA clone
IMAGE:6505924 5', mRNA sequence.

ACCESSION BU511164
VERSION BU511164.1 GI:22817397

KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.

1 (bases 1 to 1022)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. David Rowe

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM14067 Row: j column: 05

High quality sequence stop: 661.

FEATURES

source

1..1022

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:6505924"

/tissue type="undifferentiated limb"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_134"
 /notes="vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
 Cloned unidirectionally. Primer: Oligo dt. Average insert
 size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
 this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,75e-133 Length: 1022
 Score: 1392.00 Matches: 265
 Percent Similarity: 96.43% Conservatives: 5
 Best Local Similarity: 94.64% Mismatches: 4
 Query Match: 92.06% Indels: 6
 DB: 5 Gaps: 0

US-10-017-410-2 (1-275) x BU511164 (1-1022)

QY 1 MetGlyAlaProHisTyrTrpAspHisLeuArgAlaGlySerSerGluValAspTrpCys 20
 Db 120 ATGGGCCCCGCACTGGTGGACCACTCGGGCTGGCAGTTCCGAGGTGGATTTGGTGC 179
 QY 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
 Db 180 GAGGACAACTACACTATCGCTGCTGCCATTCGCCGAGTTCTACACAGATCAGCAAGTC 239
 QY 41 LeuPhePheIleLeuProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
 Db 240 TTGTTTTTCATTTTACCTCCCATCTGCATGTGCTTGTTCGCCAGTACGCAACGTGCTTC 299
 QY 61 AsnSerGlyIleTyrIleValThrLeuValValValGlyIleGlySerValTyr 80
 Db 300 AACAGCGGCATCTACTTAAATGACGCTCTCTAGTTGTAGTGGGATGGATCTGTCTAC 359
 QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpVal 100
 Db 360 TTCCATCAACGCTGAGTTTCTGGGTTCAGATGCTTGTATGACATTCCTCTCTGCTG 419
 QY 101 LeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProLysIlePheArgAsn 120
 Db 420 CTGATGTGTCTTTGGCCATGTGTTTCCAGGAGGTATTTACCAAGAGATCTTTCCGAAT 479
 QY 121 AspArgGlyValArgPheLysAlaValValCysValLeuSerAlaIleThrCysLeuAla 140
 Db 480 GACAGGGCAGGTTCAAGCAGTGTGTGTCTCTGCTGCAATTAACAAGTGTCTTGGCG 539
 QY 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
 Db 540 TTTATCAAGCCGCCATCAACAATATTTCCCTGTATGATTTCCGACATTCATGCACTGC 599
 QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
 Db 600 CTGCTTTTCAGAGCTGAAGAGGTGTGACAATGTGCGTGTGTGTTAAGCTGGGCTCTTC 659
 QY 181 SerGlyLeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysG 200
 Db 660 TCTGGCCTCTGGTGGATCTGGGCTCTTCTGTCTGGATCAGCAGCAAGCTTCTGTGA 719
 QY 200 uLeuLeuSerSerPheHisPheProTyrLeuHisCysValTrpHisIleLeuIleCysLe 220
 Db 720 GCTGCTCTCTCTCTTTCACCTTCCCTACCTGCACTGTGTGGCATATTTCTCATCTGCT 779
 QY 220 uAlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProG 240
 Db 780 TGCTTCGTACTCTGGGCTGTGTGTCTTCCCTACTTTGATGCTGCCTCAGAGATACCTGA 839
 QY 240 uGlnGlyProVal-IleArgPheTrp-ProSerGluLysTrpAlaPhe--IleGlyValP 259
 Db 840 GCCAGTCCAGTCCATCAGATTCCTGGGCCACCGAGGAATGGGCTTTTATTTGGGGTCC 899
 QY 259 roTyrValSerLeuLeuCysAla-HisLysLysSerProValLysIle 274

Db 900 CCTATGGTCCCTTCTGTGTGCCCAAGAGTCCGCGAGTCAAGATC 947

RESULT 2
 AY401891
 LOCUS Mus musculus HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence. 681 bp DNA linear GSS 12-DEC-2003

DEFINITION
 AY401891
 VERSION AY401891.1 GI:39757877
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 681)
 REFERENCE
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 681)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source 1..681
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 gene <1..>681
 /locus_tag="HCM1041"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,19e-116 Length: 681
 Score: 1230.00 Matches: 226
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 81.35% Indels: 0
 DB: 9 Gaps: 0

US-10-017-410-2 (1-275) x AY401891 (1-681)

QY 50 MetCysLeuPheArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThr 69
 Db 1 ATGTGCTTGTTCGCCAGTACGCAACGTCTTCAACAGCGCATCTACTTAATATGACG 60
 QY 70 LeuLeuValValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGly 89
 Db 61 CTCCTAGTTGTAGTGGGATTTGGATCTGTCTACTTCCATGCAACGCTGAGTTTCTCGGT 120
 QY 90 GlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPhe 109
 Db 121 CAGATGCTTGATGAACTTTCGCAATCTCTGCGGTCTGTATGTGTCTTGGCCATGTGGTT 180
 QY 110 ProArgArgTyrLeuProLysIlePheArgAsnAspArgGlyArgPheLysAlaValVal 129
 Db 181 CCCAGGAGGTATTTACCAAGATCTTTCGGAATGACAGGGGCGAGTTCGAAGGCGAGTGGT 240
 QY 130 CysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsnAsnIle 149
 Db 241 TGTGTCCTGTCTGCAATTACCAAGTGTCTGGCGTTTATCAAGCCCGCCATCAACAATATT 300
 QY 150 SerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeuLysArgCys 169

Itch, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 MEDLINE
 PUBMED
 11042159

REFERENCE

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 MEDLINE
 PUBMED
 11076861

REFERENCE

The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 20499374

REFERENCE

The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 20499374

REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayashizaki, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihata, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saigo, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

REFERENCE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.

FEATURES

URL: http://genome.gsc.riken.jp/
 Location/Qualifiers
 1..1173
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM:DB:D630008P07"
 /db_xref="taxon:10090"
 /clone="D630008P07"
 /tissue_type="kidney"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="0 day neonate"
 72..731
 /note="unnamed protein product; putative
 similar to CANCER RELATED GENE-LIVER 1 [Mus musculus]"

CDS

(SPTR|AAL40408, evidence: FASTV, 100%ID, 77.8%length,
 match=642)"
 /codon_start=1
 /protein_id="BAC39416.1"
 /db_xref="GI:26351559"
 /translation="MGAPHWMDHLRAGSGSEVDNVTYIPVAEAFYNTISNVLFFI
 LPPTCMCLPRQYATCSGIYLIWLLVVGIGSVYFATLSFLQOMLDELAILWLM
 CALAMWPPRYLPKIPENDSGRFKAVCVLSAITTCIAFLKPAINNLSIMLIGLPCTA
 LLVAELKRCNDNRVFKLGLFSLGWLWTLLALFCWISDQAFCELLSSHFHYLHCWSADR
 G"

ORIGIN

Alignment Scores:
 Pred. No.: 1e-110 Length: 1173
 Score: 1176.00 Matches: 214
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 77.78% Indels: 0
 DB: 3 Gaps: 0

US-10-017-410-2 (1-275) x AK085306 (1-1173)

Qy 1 MetGlyAlaProHisTyrTrpAspHisLeuArgAlaGlySerSerGluValAspTyrCys 20
 Db 72 ATGGCGCGCCCGCAGTCTGGTGGACCACTGGCGGCTGGCAGTTCGGAGTGGTGTGC 131
 Qy 21 GluAspAsnTyrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
 Db 132 GAGGACAACTACACTATCGTGCCTGCCATTCGCCAGTCTTACACACGATCAGCAACGTC 191
 Qy 41 LeuPhePheIleLeuProIleCysMetCysLeuPheArgGlnTyrAlaThrCysPhe 60
 Db 192 TTGTTTTTTCATTTTACCTCCCATCTGATGTGCTGTTCGCCAGTACGCAACGTCCTC 251
 Qy 61 AsnSerGlyIleTyrLeuIleTyrThrIleLeuValValValGlyIleGlySerValTyr 80
 Db 252 AACAGCGCATCTACTTAATATGACGCTCTAGTTGTAGTGGGATTTGGATCTGTCTAC 311
 Qy 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleIleuTyrVal 100
 Db 312 TTCCATGCAACGCTGAGTTTCTCGGTGAGATGCTTGTATGATGATGATGATGATGATG 371
 Qy 101 LeuMetCysAlaLeuAlaMetTyrPheProArgTyrLeuProIlePheArgAsn 120
 Db 372 CTGATGTGTGCTTTGGCCATGTGGTTTCCAGGAGGTATTTACCAAGATCTTTCGGAT 431
 Qy 121 AspArgGlyArgPheLeuAlaValValCysValLeuSerAlaIleThrThrCysLeuAla 140
 Db 432 GACAGGGGAGGTTCAAGGCGAGTGTGTGTCTCTGTCTGCAATTACCAACGCTGTGCGC 491
 Qy 141 PheIleIysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
 Db 492 TTTATCAAGCCCGCCATCAACAATATTTCCCTGTATGATCTCTGGACTTCCAGGACTCGG 551
 Qy 161 LeuLeuValAlaGluLeuLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPhe 180
 Db 552 CTGCTTGTGTGAGAGCTGAAGAGGTGTGACATGTGCGTGTGTGTTAAGCTGGGCTCTTC 611
 Qy 181 SerGlyLeuTyrTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGlu 200
 Db 612 TCTGGCTCTGTGGTGGACTCTGGCTCTCTCTGTGCTGGATCAGCAGCAACGCTTCTGTGAG 671
 Qy 201 LeuLeuSerSerPheHisPheProTyrIleHisCysValTyr 214
 Db 672 CTGCT 713

RESULT 5
 AV401890 LOCUS
 DEFINITION Pan troglodytes HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AV401890
 VERSION AV401890.1 GI:39757876

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-10-017-410-2 (1-275) x AY401890 (1-681)
QY 50 MetCysLeuPheArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThr 69
DB 1 ATGTGCTTTGTTTCGTCAGTATGCAATGCTTCAACAGTGGCATCTACTTAATCTGGACT 60
QY 70 LeuLeuValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGly 89
DB 61 CTTTGTGGTTGTAGTGGGAATGGATCCGCTACTTCCATGCAACCTTAGTTCTTGGGT 120
QY 90 GlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPhe 109
DB 121 CAGATGCTTGATGAACCTTGCACTCTTTGGGTTCGTGATGTGCTTTGGCCATGTGTTTC 180
QY 110 ProArgArgTyrLeuProIysIlePheArgAsnAspArgGlyArgPheIysAlaValVal 129
DB 181 CCCAGAAGGTATCTACCAAGATCTTTCCGAATGACCCGGGTAGGTTCGAAGTGTGGTCTC 240
QY 130 CysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsnIle 149
DB 241 ATGTGCTCTGTCGGGTATACGACATGCTGGCATTTGTCAAGCCCTGCAACATCAACATC 300
QY 150 SerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeuLysArgCys 169
DB 301 TCTCTGATGACCTGGAGTTCCTTGCACTGCTGCTATTCGAGAGCTAAGAGGTGT 360
QY 170 AspAsnValArgValPheIysLeuGlyLeuPheSerGlyLeuTrpTrpThrLeuAlaLeu 189
DB 361 GACAAATCGGTGTGTTAAGTGGTCTCTTCTCGGGCTCTGTTGGGACCTGGCCCTG 420
QY 190 PheCysTrpIleSerAspGlnAlaPheCysGluLeuSerSerPheHisPheProTyr 209
DB 421 TTCTGCTGGATCAGTACCGAGCTTTCTGCGAGCTGTGTCTATCTTCACTTCCCTAC 480
QY 210 LeuHisCysValTrpHisIleLeuIleCysLeuAlaSerTyrLeuGlyCysValCysPhe 229
DB 481 CTGCACTCATGTGGCACATCTCTATCTGCTTGTGCTGCTACCTGGCTGTGTATGCTTT 540
QY 230 AlaTyrPheAspAlaAlaSerGluIleProGluGlnGlyProValIleArgPheTrpPro 249
DB 541 GCCTACTTTGATGCTGCTCTCAGAGATTCCTGAGCAAGCCCTGTCTCATCAAGTTCTGGCC 600
QY 250 SerGluLysTyrAlaPheIleGlyValProTyrValSerLeuLeuCysAlaHisLysLys 269
DB 601 AATGAGAAATGGGCTTCTATGTTGCTCCCTATGTGTCTCTGTGTGCCAATAGAAA 660
QY 270 SerProValLysIleThr 275
DB 661 TCATCAGTCAAGATCACG 678
RESULT 6
CA463294
LOCUS
DEFINITION
IMAGE:6770649 5', mRNA sequence.
ACCESSION
CA463294
VERSION
CA463294.1 GI:24919646
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 797)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: sgabbs-remail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM3088 row: p column: 08
High quality sequence stop: 536.
FEATURES
Location/Qualifiers
1..797
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6770649"
/clone_lib="NIH MGC 169"
/note="organ: Testicles; Vector: pDNR-LIB; Site 1: Sfil
(ggcatatggcc); Site 2: Sfil (ggcgctcgcc); cDNA made
by oligo-dt priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAACGAGTGGCATTCAGCGGG-3' and
5'-ATCTAGAGCGGAGCGGCACATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
3.71e-109
Length:
Matches:
Conservative:
Mismatch:
Indels:
Gaps:
797
1158.50
7
7
76.64%
21
6

```


Db	601	AATGAGAAATGGCGCTTCATTGGTGTGCTCCCTATGTCCTCTCTGTCGCAACAGAAA	660
Qy	270	SerProVallylsileThr 275 	
Db	661	TCATCAGTCAAGATCAGC 678	
RESULT 8			
AF370405			
LOCUS	AF370405	1527 bp	linear
DEFINITION	Homo sapiens PP11646 mRNA, complete cds.		
ACCESSION	AF370405		
VERSION	AF370405.1	GI:33341739	
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1	(bases 1 to 1527)	
TITLE	Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T., Wan, D.F. and Gu, J.R.		
JOURNAL	Novel human cDNA clones with function of inhibiting cancer cell growth		
REFERENCE	Unpublished		
AUTHORS	2	(bases 1 to 1527)	
TITLE	Zhang, P.P., Zhou, X.M., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T., Wan, D.F. and Gu, J.R.		
JOURNAL	Direct Submission		
FEATURES	Submitted (16-APR-2001) National Laboratory For Oncogenes & Related Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai 200032, P. R. China		
source	Location/Qualifiers		
CDS	1..1527		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	218..640		
	/codon_start=1		
	/product="pp11646"		
	/protein_id="AAQ15241.1"		
	/db_xref="GI:33341740"		
	/translation="MCLFPROVATCFNSGIYLIWTLVVVGIGSVVFRATLSFLGOMLD ELAVLWLMCALAMWPRRYLPKIFRNDRGFKVSVLSAVTTCLAFVKPAINNISL MTLGVPTALLIAELKHEHNRNRRHKGQGGGDKV"		
ORIGIN			
Alignment Scores:			
Pred. No.:	8.7e-104	Length:	1527
Score:	1110.50	Matches:	222
Percent Similarity:	66.20%	Conservative:	13
Best Local Similarity:	62.54%	Mismatches:	4
Query Match:	73.45%	Indels:	116
DB:	3	Gaps:	2
US-10-017-410-2 (1-275) x AF370405 (1-1527)			
Qy	37	IleSerAsnValLeuPhePheIleLeuProProlleCysMetCysLeuPheArgGlnTyr 56 	
Db	179	ATCAGCAATGCTTATTATTTTCATTTTACCGCCCATCTGCATGCTGCTTTTCGTCAGTAT 238 	
Qy	57	AlaThrCysPheAsnSerGlyIleTyrLeuIleTyrThrLeuValValValGlyIle 76 	
Db	239	GCAACATGCTTCAACAGTGGCATCTACTATCTGGACTCTTTGGTTGTAGTGGAAAT 298 	
Qy	77	GlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAla 96 	
Db	299	GGATCCGCTACTTCCATGCACACCCCTAGTTCTTTGGGTCCAGATGCTTGATGAATTGCA 358 	
Qy	97	IleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgArgTyrLeuProlys 116 : :	
Db	359	GTCCTTTGGGTCTGATGTGCTTTGGCCATGTGGTTCGCCAGAGGTATCTACCAAG 418 	
Qy	117	IlePheArgAsnAspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThr 136 	
Db	419	ATCTTTCGGAATGATGACCGGGTAGGTTCAAGTGTGTGTCAGTCTCTCTCTGCGGTTACG 478 	
Qy	137	ThrCysLeuAlaPheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeu 156 	
Db	479	ACGTGCGCTGGCATTTGTCAAGCCTGCCATCAACAACATCTCTCTGATGACCCCTGGAGTT 538 	
Qy	157	ProCysThrAlaLeuLeuValAlaGluLeuLys----- 167 	
Db	539	CTTTGCACTGCACTGCTCATCGCAGAGCTAAAGAGGCATGAGAGGAACACGGAAGAGA 598 	
Qy	167	----- 167	
Db	599	CACAGGAAAGTGGCCGCAAGGAGGTGGAGACAAGGTCTCAGCATGAGTCACTCTCTGG 658 	
Qy	168	-----ArgCysAspAsnValArgValPheLysLeuGlyLeuPheSerGlyLeuTr 184 	
Db	659	ACCCCGCGTGCAGGTGTGACAACATCGGTGTGTTAAAGCTGGGCTCTTCTCGGCGCTCTG 718 	
Qy	184	pTTPThrLeuAlaLeuPheCysTTPIleSerAspGlnAlaPheCysGluLeuLeuSerSe 204 	
Db	719	GTGACCCCTGGCCCTGTTCTGCTGGATCAGTACCGAGCTTTCTCGGAGCTGCTGTCTATC 778 	
Qy	204	rPheHisPheProTyrLeuHisCysValTTP----- 214 	
Db	779	CTTCAACTTCCCTACCTGCACTGCAATGTGTGGACGGTCAGATGTTTCAGAACCACTG 838 	
Qy	214	----- 214	
Db	839	AAAGCACACTTTGTAAAAAGCTGAATTGACTCCAGGGAGCAGGCTTAGCCGGAAACGAAAG 898 	
Qy	214	----- 214	
Db	899	GCCTGATTGTCATGCTTAGCAGGACTATTAAAGAACATTTCTAAAAACGGGTCCTCAAAATTC 958 	
Qy	214	----- 214	
Db	959	TGAAGATTATCATCTCTGCTATTCCAGGGATTGTGGTAATCTTCATCTCTTCTTAATA 1018 	
Qy	215	-----HisIleLeuIleCysLeuAlaSer 222 	
Db	1019	GAAGAGGCCCTGGGCTGCAACTACAGTCAGCAAGGCACATCTCATCTGCTTGTCTGCC 1078 	
Qy	223	TyrLeuGlyCysValCysPheAlaTyrPheAspAlaAlaSerGluIleProGluGlnGly 242 	
Db	1079	TACCTGGGCTGTGTATGCTTTGCTTCTTCTGCTCTGCTCAGAGATTCCTGAGCAAGGC 1138 	
Qy	243	ProValIleArgPheTTPProSerGluLysTrpAlaPheIleGlyValProTyrValSer 262 	
Db	1139	CCTGTCAATCAAGTTCTGGCCCAATGAGAAATGGCCCTTCATTTGGTGTCCCTATGTGTTC 1198 	
Qy	263	LeuLeuCysAlaHisLysLysSerProVallylsileThr 275 	
Db	1199	CTCCTGTGTGCCAACAAAGAAATCATCAGTCAAGATCAGC 1237 	
RESULT 9			
BU234223			
LOCUS	603792086F1	698 bp	linear
DEFINITION	CSEQHN24 Gallus gallus cdna clone CHEST757k21 5', mRNA sequence.		
ACCESSION	BU234223		
VERSION	BU234223.1	GI:25478587	
KEYWORDS	EST.		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
REFERENCE	1	(bases 1 to 698)	
AUTHORS	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.		
TITLE	A Comprehensive Collection of Chicken cDNAs		
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)		
MEDLINE	22355534		

12445392
 PURNED
 COMMENT
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 FEATURES
 source
 1. 698
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, Hisex"
 /db_xref="taxon:9031"
 /clone="CHEST75K21"
 /dev_stage="22"
 /lab_host="DH10B"
 /clone_lib="CSECHN24"
 /notes="Organ: heads; Vector: pBluescript II KS(+); Site_1:
 ECORI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Alignment Scores:
 Pred. No.: 5.98e-93 Length: 698
 Score: 1001.00 Matches: 178
 Percent Similarity: 96.50% Conservativeness: 15
 Best Local Similarity: 89.00% Mismatches: 7
 Query Match: 66.20% Indels: 0
 DB: 5 Gaps: 0

US-10-017-410-2 (1-275) x BU234223 (1-698)

QY 76 IleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeu 95
 Db 1 ATTGGATCTCTTTACTTTTCATGTCACCCCTCAGCTTCTGGGTTCAGATGCTGAGCTG 60

QY 96 AlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyrLeuPro 115
 Db 61 GCATTCCTGGGTCTTATGTCCTTTGGCATGGTTCCTTAGGATATCTACCA 120

QY 116 LysIlePheArgAsnAspArgGlyArgPheLysAlaValValCysValLeuSerAlaIle 135
 Db 121 AGAGTTTTTCGAAATGACAGAGCGCGTTTAAAGCTGCTGTGTGTCCTGCTGGAGTT 180

QY 136 ThrThrCysLeuAlaPheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGly 155
 Db 181 ACTACCTGCTTTCCTTTCATTAACCTGCCATCAACCACTCACTAAATGACTCTGGGT 240

QY 156 LeuProCysThrAlaLeuValAlaGluLeuLysArgCysAspAsnValArgValPhe 175
 Db 241 GTTCCTTGCACAGCTTTACTCATGCTGAGTTGAAGAGGTGTGAACACCTGGGTGTAC 300

QY 176 LysLeuGlyLeuPheSerGlyLeuTrpTrpTrpTrpLeuAlaLeuPheCysTrpIleSerAsp 195
 Db 301 AAGCTTGCTGCTGTTTTCAGGTCTTTGGTGATGCTAGCACTTTTCTGCTGGATCAGTGAC 360

QY 196 GlnAlaPheCysGluLeuLeuSerPheHisPheProTyrLeuHisCysValTrpHis 215
 Db 361 AAAGCTTTTGTGAGATCTGTGTCATCACTTAACTTCCCTTATTGCACTGTGTATGGCAC 420

QY 216 IleLeuIleCysLeuAlaSerTyrLeuGlyCysValCysPheAlaTyrPheAspAlaAla 235
 Db 421 ATTTTGAATTTGCTTTCGGCTTACCTAGGCTGTGTCTGTTCGCTTACTTCGATGCTGCC 480

QY 236 SerGluIleProGluGlnGlyProValIleArgPheTrpProSerGluLysTrpAlaPhe 255
 Db 481 TCCAGATCCCTGAGCAGGCGCCCGTCATTAAGTTCTGGCCAGTAGAGATGGGCATTC 540

QY 256 IleGlyValProTyrValSerLeuLeuCysAlaHisLysLysSerProValLysIleThr 275
 Db 541 ATTGGCGTTCCTTACGTCACCTCTCTGTGTCACACAAAGAAATCACCGGTGAAGATCACA 600

RESULT 10
 CF115220 565 bp mRNA linear EST 23-JUL-2003
 LOCUS Shultzomica08471 Rat lung airway and parenchyma cDNA libraries
 DEFINITION Rattus norvegicus cDNA clone NP6159 5', mRNA sequence.
 ACCESSION CF115220
 VERSION CF115220.1 GI:33175919
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 565)
 AUTHORS Shultz, M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V.,
 Padua, A.M., Gurske, W.A., Morin, D., Penn, S.G., Jovanovich, S.B.,
 Plopper, C.G. and Buckpitt, A.R.
 TITLE Gene expression analysis in response to lung toxicants: I.
 Sequencing and microarray development
 JOURNAL Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
 COMMENT Contact: Shultz MA
 Dept. of Molecular Biosciences, School of Veterinary Medicine
 University of California, Davis
 1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
 Tel: 530 752 0793
 Fax: 530 752 4698
 Email: mashultz@ucdavis.edu
 Average Phred score is 20 or better. All poor quality data (Phred <
 20) and vector/linker sequence has been removed.
 High quality sequence stop: 565.

FEATURES
 Location/Qualifiers
 1..565
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="NP6159"
 /sex="male"
 /tissue_type="airway or parenchyma"
 /dev_stage="adult"
 /clone_lib="Rat lung airway and parenchyma cDNA libraries"
 /notes="Organ: lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
 Site 2: Not I; mRNA was isolated from microdissected rat
 lung airways and parenchyma tissues."

ORIGIN

Alignment Scores:
 Pred. No.: 1.45e-92 Length: 565
 Score: 996.00 Matches: 181
 Percent Similarity: 99.47% Conservativeness: 6
 Best Local Similarity: 96.28% Mismatches: 0
 Query Match: 65.87% Indels: 1
 DB: 6 Gaps: 0

US-10-017-410-2 (1-275) x CF115220 (1-565)

QY 74 ValGlyIleGlySerValTyrPheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAs 93
 Db 3 GTGGGAATCGAGTCTGTCTACTTCCATNGCAACTCTTAGTTCTCTGGTCAGATGCTTGA 62

QY 93 pGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyr 113

Db 309 TTCCATGCAACGCTGAGTTTCTCTGGGTGAGATGCTTGATGAAGCTTGCCATTTCTGTGGGTT 368
 QY 101 LeuMetCysAlaLeuAlaMetTrrPheProArgTyrLeuProLysIlePheArgAsn 120
 Db 369 CTGATGTGTCTTTGGCCATGTGGTTTCCAGGAGGTATTACCAAGATCTTTCCGAAT 428
 QY 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThrThrCysLeuAla 140
 Db 429 GACAGGGCAGGTTCAAGGAGGTGTGTCTCTGCTCAATTACAGGTGTGTGGC 488
 QY 141 PheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAla 160
 Db 489 TTTATCAAGCCGCCATCAACATATTTCCCTGATGATTTCTGGACTTCCATGCACTGGC 548
 QY 161 LeuLeuValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGly-LeuPh 180
 Db 549 CTGCTTGTTCAGAGCTGAAAGGTGTGACAATGTGGCTGTGTGTAAATGGGCCCTTTT 608
 QY 180 eSerGlyLeuTrrPheThrLeuAla 188
 Db 609 TTTGGCCTTGTGGTGGACTCTGGCT 633

RESULT 12

CB723138

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB723138 622 bp mRNA linear EST 09-JUL-2003
 UI-M-GHO-ceh-f-04-0-UI.r1 NIH BMAP_GHO Mus musculus cDNA clone
 IMAGE:6839525 5', mRNA sequence.
 CB723138.1 GI:29780280
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 622)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strauberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers

1..622
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6839525"
 /tissue_type="Whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP_GHO"
 /note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CGAAGTGAAT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN
 Alignment Scores: 2.9e-91 Length: 622
 Pred. No.: 984.00 Matches: 180
 Score: 78.95% Conservative: 0
 Percent Similarity: 78.95% Mismatches: 2
 Best Local Similarity: 78.95% Indels: 46
 Query Match: 65.08% Gaps: 1
 DB: 6
 US-10-017-410-2 (1-275) x CB723138 (1-622)
 QY 1 MetGlyAlaProHisTrrPheAspHisLeuArgAlaGlySerSerGluValAspTrrPheCys 20
 Db 76 ATGGGGCCCGCCACTGGTGGGACCCCTCGGGCTGGCAGTTCGGAGGTGGATTGTGTC 135
 QY 21 GluAspAsnTrrThrIleValProAlaIleAlaGluPheTyrAsnThrIleSerAsnVal 40
 Db 136 GAGGACNACTACACTATCGTCCCTGCAATGCGAGTTCTACAACAGATCAGCAACGTC 195
 QY 41 LeuPhePheIleLeuProProIleCysMetCysLeuPheArgGlnTrrAlaThrCysPhe 60
 Db 196 TTGTTTTTCATTTTACCTCCCATCTGCATGTGCTTGTTCGCCAGTACGCAACGTCGCTTC 255
 QY 61 AsnSerGlyIleTrrLeuIleTrrPheLeuValValGlyIleGlySerValTyr 80
 Db 256 AACAGCGGCATCTACTTAATATGAGCCTCTAGTTAGTGGGATTTGGATCTGTCTAC 315
 QY 81 PheHisAlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrrPheVal 100
 Db 316 TTCCATGCAACGCTGAGTTTCTGGGTGACATGCTTGATGAAGTTCGCTTCTGTGGGTT 375
 QY 101 LeuMetCysAlaLeuAlaMetTrrPheProArgTyrLeuProLysIlePheArgAsn 120
 Db 376 CTGATGTGTCTTTGGCCATGTGGTTTCCAGGAGGTATTACCAAGATCTTTCCGAAT 435
 QY 121 AspArgGlyArgPheLysAlaValValCysValLeuSerAlaIleThrThrCysLeuAla 140
 Db 436 GAC-----AGGTGTGACAATGTGCTGTGTTAAAGCTGGGCTCTCTTC 477
 QY 181 SerGlyLeuTrrPrrThrLeuAlaLeuPheCysTrrPheSerAspGlnAlaPheCysGlu 200
 Db 478 TCTGGCCTCTGGTGGACTCTGGCTCTCTTCTGTGGATCAGCCAGCCATCTCTGTGAG 537
 QY 201 LeuLeuSerSerPheHisPheProTrrLeuHisCysValTrrPheIleLeuIleCysLeu 220
 Db 538 CTGCTCTCTCTCTTTCATTTCCCTACTGCACTGTGTGGCATATTTCTCATCTGCTT 597
 QY 221 AlaSerTrrLeuGlyCysValCys 228
 Db 598 GCTTCGTACTGGGGCTGTGTGTGC 621
 RESULT 13
 BU232554
 LOCUS
 DEFINITION
 603409105F1 CSBQCHN24 Gallus gallus cdna clone CHEST323j2 5', mRNA
 sequence.
 BU232554
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Gallus gallus (chicken)
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

1 (bases 1 to 781)

REFERENCE AUTHORS

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

12445392

Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

FEATURES

source

1. 781
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST323j2"
/dev_stage="22"
/lab_host="DH10B"
/clone_lib="CSEQCEN24"
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following the first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

RESULT 14
CF169369
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
COMMENT

CF169369 548 bp mRNA linear EST 25-JUL-2003
B0812G07-5 NIA Mouse Newborn Kidney cDNA Library (long 1) Mus
musculus cDNA clone NIA:B0812G07 IMAGE:30468654 5', mRNA sequence.
CF169369
CF169369.1 GI:33278918
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 548)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
21429098
11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0812 row: G column: 07
Seq primer: M13 Reverse
High quality sequence stop: 548
POLYA=No.

US-10-017-410-2 (1-275) x BU232554 (1-781)

Alignment Scores:
Pred. No.: 4.72e-89 Length: 781
Score: 964.00 Matches: 177
Percent Similarity: 95.05% Conservative: 15
Best Local Similarity: 87.62% Mismatches: 8
Query Match: 63.76% Indels: 2
DB: 5 Gaps: 0

US-10-017-410-2 (1-275) x BU232554 (1-781)

QY 76 lIeGlySerValTyrPheHis-AlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLe 95
Db 1 ATTGGATCTGTTTACTTTCATGCCACCCCTCAGCTTCCTGGGTGAGATGCTGAGTACT 60
QY 95 uAlaIleLeuTrpValLeuMetCysAlaLeuAlaMetTrpPheProArgTyrTyrLeuPr 115
Db 61 GGCATTCTCTGGTGCCTTATGTGTGCTTGTGCCATGTGTTCCTCCTAGGAGATATCTACC 120
QY 115 oLysIlePheArgAnAspArgGlyArgPheLysAlaValValCysValLeuSerAlaI 135
Db 121 AAGAGTTTTTCGAATGACAGAGCGGTTTAAAGCTGCTGTGTGCTGCTGTGGAGT 180
QY 135 eThrThrCys-LeuAlaPheIleLysProAlaIleAsnAsnIleSerLeuMetIleLeuG 155
Db 181 TACTACCTGGCGCTTCCTTCATTAAACCTGCCATCAACAACTCTCATTAAAGTACTTGG 240
QY 155 lYLeuProCysThrAlaLeuLeuValAlaGluLeuLysArgCysAspAnValArgValP 175
Db 241 GTGTTTCCTTGCACAGCTTACTCATGTGCTGAGTTCGAGAGGTGTGAACCTGCGTGTGT 300

FEATURES

source

Location/Qualifiers
1. 548
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0812G07-5"
/db_xref="taxon:10090"
/clone="NIA:B0812G07 IMAGE:30468654"
/dev_stage="Newborn Kidney"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
In brief, double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen:
5'-PGATAGTTCATGATCGGAGCGGCCCTTTTCTTTT-3'] from
26 ug of total RNA, treated with 14 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0 kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
 Pred. No.: 1.01e-84 Length: 548
 Score: 920.00 Matches: 179
 Percent Similarity: 78.95% Conservative: 1
 Best Local Similarity: 78.51% Mismatches: 2
 Query Match: 60.85% Indels: 46
 DB: 6 Gaps: 1

US-10-017-410-2 (1-275) x CF169369 (1-548)

```

QY 8 AspHisLeuArgAlaGlySerSerGluValAspTrpCysGluAspAsnTyrThrIleVal 27
Db 1 GACCACCTGGGAGTGGAGTTCGGAGTGGATGGTGGAGGACAACTACACTATCGTG 60
QY 28 ProAlaIleAlaGluPheTyrAsnThrIleSerAsnValLeuPhePheIleLeuProPro 47
Db 61 CCTGCCATTGCCAGTCTCAACACGATCAGCAACGCTCTGTGTTTCAITTTACCTCCC 120
QY 48 IleCysMetCysLeuPheArgGlnTyrAlaThrCysPheAsnSerGlyIleTyrLeuIle 67
Db 121 ATCTGCATGCTGTGTTCCGCCAGTAGTCGCAACGCTTCAACGCGCATCTACTTAATA 180
QY 68 TrpThrLeuLeuValValGlyIleGlySerValTyrPheHisAlaThrLeuSerPhe 87
Db 181 TGGACGCTCTAGTCTAGTGGGATGGATCTGTCTACTTCCATGCAACGCTGAGTTTC 240
QY 88 LeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMetCysAlaLeuAlaMet 107
Db 241 CTGGGTGAGTGGCTGTGCAATTACAACTGCTTGGGTTTATCAAGCCGCGCATCAAC 420
QY 108 TrpPheProArgArgTyrLeuProGlyIlePheArgAsnAspArgGlyArgPheLysAla 127
Db 301 TGGTTTCCCGAGGAGTATTTACCAAGATCTTTCCGGATGACAGGGGCGAGTCAAGGCA 360
QY 128 ValValCysValLeuSerAlaIleThrThrCysLeuAlaPheIleLysProAlaIleAsn 147
Db 361 GTGGTGTGGCTGTGCTGCAATTACAACTGCTTGGGTTTATCAAGCCGCGCATCAAC 420
QY 148 AsnIleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeuValAlaGluLeuLys 167
Db 421 AATATTTCCTGATGATCTGGGACTTCCATGACCTCGCTGCTTGTGTCAGAGCTGAAG 480
QY 168 AtqCysAspAsnValargValPheLysLeuGlyLeuPheSerGlyLeuTrpThrLeu 187
Db 481 AGG----- 483
QY 188 AlaLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeuSerPheHisPhe 207
Db 483 ----- 483
QY 208 ProTyrLeuHisCysValTrpHisIleLeuIleCysLeuAlaSerTyrLeuGlyCysVal 227
Db 484 -----CATATTCTCATCTGCTTCTCGTACCTGGGCTGTGTG 522
QY 228 CysPheAlaTyrPheAspAlaAla 235
Db 523 TGGTTGCGCTACTTTGATGCTGCC 546

```

RESULT 15

BJ060795

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BJ060795 648 bp mRNA linear EST 29-SEP-2003
 BJ060795 NIBB Mochii normalized Xenopus tailbud library Xenopus
 laevis cDNA clone XL066n16 5', mRNA sequence.
 BJ060795
 BJ060795.1 GI:17500536
 EST.
 Xenopus laevis (African clawed frog)
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 648)
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and
 Kohara, Y.
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 Contact: Tadashi Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp
 The information of this clone is available through the following
 URL.
 http://xenopus.nibb.ac.jp.

FEATURES

source

Location/Qualifiers

1..648

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="XL066n16"

/tissue_type="whole embryo"

/dev_stage="stage 25"

/clone_lib="NIBB Mochii normalized Xenopus tailbud
 library"

ORIGIN

Alignment Scores:

Pred. No.: 2.23e-83 Length: 648

Score: 908.00 Matches: 162

Percent Similarity: 93.78% Conservative: 19

Best Local Similarity: 83.94% Mismatches: 12

Query Match: 60.05% Indels: 0

DB: 4 Gaps: 0

US-10-017-410-2 (1-275) x BJ060795 (1-648)

QY 83 AlaThrLeuSerPheLeuGlyGlnMetLeuAspGluLeuAlaIleLeuTrpValLeuMet 102

Db 2 GCACCCCTCAGCTTTCTCGGGCAGATGCTGGATGAGTTGGCTATCTCTGGGTGCTTATG 61

QY 103 CysAlaLeuAlaMetTrpPheProArgTyrLeuProLysIlePheArgAsnAspArg 122

Db 62 TCTGCGTGGCCATGTGGTTCCCTAAAGACACTTGCACAGGGTATTTCCGAATGACAGG 121

QY 123 GlyArgPheLysAlaValValCysValLeuSerAlaIleThrThrCysLeuAlaPheIle 142

Db 122 TGGAGGTTTAAAGCAGTTGTTGGGGTGAATTCGGGAGTCCACACAGCCCTAGCGTTTATC 181

QY 143 LysProAlaIleAsnAsnIleSerLeuMetIleLeuGlyLeuProCysThrAlaLeuLeu 162

Db 182 AACCTTCCCATTAACAGCATATCTCTGATGATTTCTTGGGATACCTTGTAGCTTACTC 241

QY 163 ValAlaGluLeuLysArgCysAspAsnValArgValPheLysLeuGlyLeuPheSerGly 182

Db 242 ATTGCTGAGCTCAAGAGATGTGACATATATGCGGGTTTAAAGCTTGGAGTGATGCTGGG 301

QY 183 LeuTrpTrpThrLeuAlaLeuPheCysTrpIleSerAspGlnAlaPheCysGluLeuLeu 202

Db 302 TTTTGGTGGACGCTGGCGTGTGGCTCTGTGGATCAGTGACAAAGCTTTTGTGAATATGG 361

This Page Blank (uspto)

construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-f. 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

source 1..2852

Location/Qualifiers

1..2852

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CTONG202638"

/tissue type="tongue, tumor tissue"

/clone_lib="CTONG2"

/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 98.1%; Score 812; DB 9; Length 2852;

Best Local Similarity 98.8%; Pred. No. 8.9e-210;

Matches 818; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCAGCTGGTGGGACGAGCTGAGGCTGGTGGAGTGGAGTGGTGC 60

Db 159 ATGGGCGCCCGCAGCTGGTGGGACGAGCTGAGGCTGGTGGAGTGGAGTGGTGC 218

QY 61 GAGGACAACTACACCATCGCTGCTGCTATCGCCGAGTCTTACACAGATCAGCAATGTC 120

Db 219 GAGGACAACTACACCATCGCTGCTGCTATCGCCGAGTCTTACACAGATCAGCAATGTC 278

QY 121 TTAATTTTCAATTTTACCGCCCATCTGCATGTGCTTGTGATGAGTATGCAATGCTTG 180

Db 279 TTAATTTTCAATTTTACCGCCCATCTGCATGTGCTTGTGATGAGTATGCAATGCTTG 338

QY 181 AACAGTGCATCTACTTAATCTGGACTCTTTGGTGTGAGTGGAAATGGATCGCTTAC 240

Db 339 AACAGTGCATCTACTTAATCTGGACTCTTTGGTGTGAGTGGAAATGGATCGCTTAC 398

QY 241 TTCCATTTTACCCCTTAGTCTTGGGTGAGTGTGATGATGAACTTGCAGTCTTTTGGGT 300

Db 399 TTCCATTTTACCCCTTAGTCTTGGGTGAGTGTGATGATGAACTTGCAGTCTTTTGGGT 458

QY 301 CTGATGTGTCTTTGGGCAATGTGTTCCCGAAGAGTATCTTACCAAGATCTTTGGGAT 360

Db 459 CTGATGTGTCTTTGGGCAATGTGTTCCCGAAGAGTATCTTACCAAGATCTTTGGGAT 518

QY 361 GACAGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 420

Db 519 GACAGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 578

QY 421 TTTGTCAGGCTGCGCATCAACAACTCTCTGATGACCCCTGGGAGTTCCTTGGCACTGCA 480

Db 579 TTTGTCAGGCTGCGCATCAACAACTCTCTGATGACCCCTGGGAGTTCCTTGGCACTGCA 638

QY 481 CTGCTCATTCGAGAGCTAAAGAGTGTGACAAATGCGTGTGTTAAGCTGGGCTCTTC 540

Db 639 CTGCTCATTCGAGAGCTAAAGAGTGTGACAAATGCGTGTGTTAAGCTGGGCTCTTC 698

QY 541 TCGGGCTCTGTTGGTGGACCTGGCCCTGTTCTGCTGAGTGTGAGTGTGAGTGTGAGT 600

Db 699 TCGGGCTCTGTTGGTGGACCTGGCCCTGTTCTGCTGAGTGTGAGTGTGAGTGTGAGT 758

QY 601 CTGCTGTCTCATCTTCAACTTCCCTCTCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 660

Db 759 CTGCTGTCTCATCTTCAACTTCCCTCTCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 818

QY 661 GCTGCTACCTGGGCTGTGATGCTTTCCTTACTTGTGATGCTGCTGAGATTCCTGAG 720

Db 819 GCTGCTACCTGGGCTGTGATGCTTTCCTTACTTGTGATGCTGCTGAGATTCCTGAG 878

QY 721 CAAGCCCTGTGATCAAGTCTTGGCCCAATGAGAAATCATCAAGTCAAGATCAAGTGA 780

Db 879 CAAGCCCTGTGATCAAGTCTTGGCCCAATGAGAAATCATCAAGTCAAGATCAAGTGA 938

QY 781 GTGTCCTCTCTGTTGGTGGCAACAGAAATCATCAAGTCAAGATCAAGTGA 828

Db 939 GTGTCCTCTCTGTTGGTGGCAACAGAAATCATCAAGTCAAGATCAAGTGA 986

339 AACAGTGGCATCTACTTAATCTGGACCTCTTTTGGTGTGAGTGGAAATGGATCGGTCTAC 398

241 TTCCATTTTACCCCTTAGTCTTGGGTGAGTGTGATGATGAACTTGCAGTCTTTGGGTT 300

399 TTCCATGCAACCTTAGTCTTGGGTGAGTGTGATGATGAACTTGCAGTCTTTGGGTT 458

301 CTGATGTGTCTTTGGGCAATGTGTTCCCGAAGAGTATCTTACCAAGATCTTTGGGAT 360

459 CTGATGTGTCTTTGGGCAATGTGTTCCCGAAGAGTATCTTACCAAGATCTTTGGGAT 518

361 GACAGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 420

519 GACAGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 578

421 TTTGTCAGGCTGCGCATCAACAACTCTCTGATGACCCCTGGGAGTTCCTTGGCACTGCA 480

579 TTTGTCAGGCTGCGCATCAACAACTCTCTGATGACCCCTGGGAGTTCCTTGGCACTGCA 638

481 CTGCTCATTCGAGAGCTAAAGAGTGTGACAAATGCGTGTGTTAAGCTGGGCTCTTC 540

639 CTGCTCATTCGAGAGCTAAAGAGTGTGACAAATGCGTGTGTTAAGCTGGGCTCTTC 698

541 TCGGGCTCTGTTGGTGGACCTGGCCCTGTTCTGCTGAGTGTGAGTGTGAGTGTGAGT 600

699 TCGGGCTCTGTTGGTGGACCTGGCCCTGTTCTGCTGAGTGTGAGTGTGAGTGTGAGT 758

601 CTGCTGTCTCATCTTCAACTTCCCTCTCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 660

759 CTGCTGTCTCATCTTCAACTTCCCTCTCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 818

661 GCTGCTACCTGGGCTGTGATGCTTTCCTTACTTGTGATGCTGCTGAGATTCCTGAG 720

819 GCTGCTACCTGGGCTGTGATGCTTTCCTTACTTGTGATGCTGCTGAGATTCCTGAG 878

721 CAAGCCCTGTGATCAAGTCTTGGCCCAATGAGAAATCATCAAGTCAAGATCAAGTGA 780

879 CAAGCCCTGTGATCAAGTCTTGGCCCAATGAGAAATCATCAAGTCAAGATCAAGTGA 938

781 GTGTCCTCTCTGTTGGTGGCAACAGAAATCATCAAGTCAAGATCAAGTGA 828

939 GTGTCCTCTCTGTTGGTGGCAACAGAAATCATCAAGTCAAGATCAAGTGA 986

RESULT 2

AK123581

LOCUS 2852 bp mRNA linear PRI 09-SEP-2003

DEFINITION Homo sapiens cDNA FLJ41587 f1e, clone CTONG2020638.

ACCESSION AK123581

VERSION AK123581.1 GI:34529163

KEYWORDS oligo capping; f1s (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

2 Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuo, Y., Hagi, K. and Isogai, T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE

1 (bases 1 to 2852)

2 Isogai, T. and Yamamoto, J.

DIRECT SUBMISSION

SUBMITTED (15-JUN-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kibarasu, Chiba 292-0818, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library

Db 939 GTGTCCTCTCTGTGTCACCAAGAAATCATCAGTCAAGATCACGTGA 986
|||||

RESULT 3

AC017081/c
LOCUS AC017081 149462 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-470J24 from 2, complete sequence.
ACCESSION AC017081
VERSION AC017081.8 GI:18093316
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Nguyen,C., Doebber,A. and Kozlowski,A.

TITLE The sequence of Homo sapiens BAC clone RP11-470J24

JOURNAL Unpublished (2001)

REFERENCE

AUTHORS Waterston,R.H.

TITLE Direct Submission

JOURNAL Submitted (09-DEC-1999) Genome Sequencing Center, Washington

MO 63108, USA

4 (bases 1 to 149462)

Waterston,R.

Direct Submission

Submitted (09-JAN-2002) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Jan 9, 2002 this sequence version replaced gi:14165368.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc

Contact: sapiens@watson.wustl.edu

----- Summary Statistics

----- Center project name: H_NH0470J24

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping

clone sections once, or longer because we provide a small overlap

between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by

restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.

McPherson, Department of Genetics, Washington University, St. Louis

MO. For additional information about the map position of this

sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male

donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E.,

Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved

approach for construction of bacterial artificial chromosome

libraries. Genomics 51:1-8. The clone may be obtained either from

Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong

and coworkers at the Roswell Park Cancer Institute

(http://bacpac.med.buffalo.edu)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-310K15, 2000 bp overlap;

the clone sequenced to the right is RP11-156A1, 2000 bp overlap.

Actual start of this clone is at base position 190775 of

RP11-310K15; actual end is at base position 28935 of RP11-156A1.

The sequence between 81601 and 81703 is covered only by a PCR
product of clone DNA. Unresolved tandem repeat regions exist
between 81548 and 83183, 126196 and 127426.

FEATURES

source

1. .149462

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-470J24"

/clone_lib="RPC1-11"

141. .282

/rpt_family="MIR"

581. .621

/rpt_family="(TGGG)n"

965. .1037

/rpt_family="GC-rich"

1127. .1561

/note="match to EST AA056210 (NID:g1548612) zf62g06.s1"

1771. .2127

/note="match to EST AW614820 (NID:g7320006) hg80c08.x1"

2055. .2195

/rpt_family="MER1_type"

2216. .2419

/rpt_family="MER1_type"

2604. .2844

/rpt_family="MIR"

2991. .3324

/note="match to EST R35058 (NID:g791959) yh86d09.r1"

3207. .3758

/note="match to EST AW452004 (NID:g6992780) "

3486. .3527

/rpt_family="AT-rich"

3501. .3745

/note="match to EST AA054208 (NID:g1545131) zf51g03.r1"

3981. .4017

/rpt_family="T-rich"

4231. .4273

/rpt_family="AT-rich"

5518. .5847

/note="match to EST R80994 (NID:g857275) y194c08.r1"

5606. .5633

/rpt_family="AT-rich"

5722. .5758

/rpt_family="AT-rich"

5802. .5831

/rpt_family="(TTTTA)n"

5808. .6116

/rpt_family="Alu"

6153. .6244

/rpt_family="MER1_type"

6965. .7000

/rpt_family="AT-rich"

7020. .7094

/rpt_family="MER1_type"

7982. .8006

/rpt_family="AT-rich"

8000. .8436

/rpt_family="L2"

8437. .8540

/rpt_family="L1"

8541. .8696

/rpt_family="L2"

8722. .9684

repeat_region

Best Local Similarity 96.0%; Pred. No. 6.2e-196; Matches 792; Conservative 0; Mismatches 30; Indels 3; Gaps 1;									
QY	1	ATGGCGCCCGC	ACTGTTGGAC	CAGCTG	CGAGCTG	TGAGT	CGAGT	GGTGG	ACTGGTGC
Db	1	ATGGCGCCCGC	ACTGTTGGAC	CAGCTG	CGAGCTG	CGAGCTG	CGAGCTG	CGAGCTG	CGAGCTG
QY	61	GAGGCAACTAC	CACATCGT	CGCTG	TATCG	CGGAG	TTCTA	CAACAC	GCATGATGTC
Db	61	GAGGCAACTAC	CACATCGT	CGCTG	TATCG	CGGAG	TTCTA	CAACAC	GCATGATGTC
QY	121	TTATTTTTCAT	TTTACCG	CCCATCT	GCATG	CTCTG	TTTGT	GATG	ATGCAACATGCTTG
Db	121	TTATTTTTCAT	TTTACCG	CCCATCT	GCATG	CTCTG	TTTGT	GATG	ATGCAACATGCTTG
QY	181	AACAGTGACAT	CTACTTAAT	CTGGACT	CTTTTGG	TGTAG	TGGAA	TGGAT	TGGATCTTAC
Db	181	AACAGTGACAT	CTACTTAAT	CTGGACT	CTTTTGG	TGTAG	TGGAA	TGGAT	TGGATCTTAC
QY	241	TTCCATTTTAC	CTTGTAG	TTCTTGG	TGATG	CTTGT	GATG	AACTT	GCAGTCTTGGGTT
Db	241	TTCCATTTTAC	CTTGTAG	TTCTTGG	TGATG	CTTGT	GATG	AACTT	GCAGTCTTGGGTT
QY	301	CTGATGTGCT	TTTGG	CCATG	TTTCC	CAGAG	STAT	CTAC	CAAGATCTTTCGGAT
Db	301	CTGATGTGCT	TTTGG	CCATG	TTTCC	CAGAG	STAT	CTAC	CAAGATCTTTCGGAT
QY	361	GACAGGGTAG	TTTCAAG	TGGTGG	TGCTG	CTG	CGGTT	TAC	GACGCTGGCA
Db	361	GACAGGGTAG	TTTCAAG	TGGTGG	TGCTG	CTG	CGGTT	TAC	GACGCTGGCA
QY	421	TTTGTCAAG	CCCTGC	ATCA	CAACAT	CTCT	CTG	ATG	ACCTGGAGTCTTGGCACTGCA
Db	421	TTTGTCAAG	CCCTGC	ATCA	CAACAT	CTCT	CTG	ATG	ACCTGGAGTCTTGGCACTGCA
QY	481	CTGCTCAT	CGCAG	CTTAA	GAGG	TGTG	CAACAT	CGT	CTTAAAGCTGGGCTCTTTC
Db	481	CTGCTCAT	CGCAG	CTTAA	GAGG	TGTG	CAACAT	CGT	CTTAAAGCTGGGCTCTTTC
QY	541	TCGGGCTCT	GTGGAC	CCCTG	GGCTG	TCTG	CTG	ATG	ACCGAGCTTTCTGCGAG
Db	541	TCGGGCTCT	GTGGAC	CCCTG	GGCTG	TCTG	CTG	ATG	ACCGAGCTTTCTGCGAG
QY	601	CTGCTGTAT	CTTCA	ACTTCC	CTAC	TGCA	CTG	CA	TCATCTATGCTCTT
Db	601	CTGCTGTAT	CTTCA	ACTTCC	CTAC	TGCA	CTG	CA	TCATCTATGCTCTT
QY	661	GCTGCTAC	CTGGG	CTGTAT	GCTTT	GGCTACT	TTGAT	GCCT	CTCAGAGATCTCTGAG
Db	661	GCTGCTAC	CTGGG	CTGTAT	GCTTT	GGCTACT	TTGAT	GCCT	CTCAGAGATCTCTGAG
QY	721	CAAGCCCTGT	CATCA	AGTTCT	TGGCC	CAATG	AGAA	TGG	CCCTTCATTGGTGTCCCTTAT
Db	721	CAAGCCCTGT	CATCA	AGTTCT	TGGCC	CAATG	AGAA	TGG	CCCTTTCATTGGTGTCCCTTAT
QY	781	GTGTCCTCT	GTGTG	CCCA	CAAG	AAATCAT	CAGT	CA	AGTCAAG
Db	781	GTGTCCTCT	GTGTG	CCCA	CAAG	AAATCAT	CAGT	CA	AGTCAAG
RESULT 5									
AF282864									
LOCUS AF282864									
DEFINITION Mus musculus cancer related gene-liver 1 mRNA, complete cds.									
ACCESSION AF282864									
VERSION AF282864.1									
KEYWORDS GI:17529683									
SOURCE Mus musculus (house mouse)									
ORGANISM Mus musculus									
REFERENCE 1 (bases 1 to 4174)									
AUTHORS Graveel, C.R., Jackoe, T., Madore, S.J., Holt, A.L. and Farnham, P.J.									
TITLE Expression profiling and identification of novel genes in hepatocellular carcinomas									
JOURNAL Oncogene 20 (21), 2704-2712 (2001)									
MEDLINE 21313787									
PUBMED 11420682									
REFERENCE 2 (bases 1 to 4174)									
AUTHORS Graveel, C.R., Jackoe, T., Madore, S.J., Holt, A.L. and Farnham, P.J.									
TITLE Identification of genes deregulated in murine hepatocellular carcinomas using oligonucleotide microarrays and representational difference analysis									
JOURNAL Unpublished									
REFERENCE 3 (bases 1 to 4174)									
AUTHORS Farnham, P.J. and Graveel, C.R.									
TITLE Direct Submission									
JOURNAL Submitted (27-JUN-2000) Oncology, University of Wisconsin, 1400									
FEATURES Location/Qualifiers									
1..4174									
/organism="Mus musculus"									
/mol_type="mRNA"									
/db_xref="taxon:10090"									
/tissue_type="hepatocellular carcinoma"									
35..862									
/note="CRG-L1; putative"									
/codon_start=1									
/product="cancer related gene-liver 1"									
/protein_id="AAL40408.1"									
/db_xref="GI:17529684"									
/translation="MGAPMWDHLRAGSEVDWEDNYTIVPAIEFYNTISNVLFFI									
LPPICMLFRQYATCFNSGIYLIWLLVVGIGSVYFPHATLSFLGMDLDELAILVLM									
CALAMFPRYLPKIFIRDRGRFRAVCVLSAITLAFIPAINNISLGLPCTA									
LLVLAELKRCNDNRVFKLGFSLGMLWTLFCWISDQACELLSFHFYLLHCVHILI									
CLASYLGVCVFAYFDAASEIPEQGPVIRFMPSEKWFATIGVPYVLSLLCAHKKSVPKIT"									
ORIGIN									
Query Match 80.3%; Score 664.8; DB 10; Length 4174;									
Best Local Similarity 87.7%; Pred. No. 1.1e-169;									
Matches 726; Conservative 0; Mismatches 102; Indels 0; Gaps 0;									
QY	1	ATGGCGCCCGC	CACTGGT	GGAGC	CCAGCT	GAGGCT	GAGT	CGGAGT	GGTGGTGC
Db	35	ATGGCGCCCGC	CACTGGT	GGAGC	CCAGCT	GAGGCT	GAGT	CGGAGT	GGTGGTGC
QY	61	GAGGCAACTAC	ACATCGT	CGCTG	CTATG	CGCCG	AGTTCT	ACAA	CGATCAGCAATGTC
Db	95	GAGGCAACTAC	ACATCGT	CGCTG	CTATG	CGCCG	AGTTCT	ACAA	CGATCAGCAATGTC
QY	121	TTATTTTTCAT	TTTACCG	CCCATCT	GCATG	CTCTG	TTTGT	GATG	ATGCAACATGCTTG
Db	155	TTATTTTTCAT	TTTACCG	CCCATCT	GCATG	CTCTG	TTTGT	GATG	ATGCAACATGCTTG
QY	181	AACAGTGACAT	CTACTTAAT	CTGGACT	CTTTTGG	TGTAG	TGGAA	TGGAT	TGGATCTTTCGGAT
Db	215	AACAGTGACAT	CTACTTAAT	CTGGACT	CTTTTGG	TGTAG	TGGAA	TGGAT	TGGATCTTTCGGAT
QY	241	TTCCATTTTAC	CTTGTAG	TTCTTGG	TGATG	CTTGT	GATG	AACTT	GCAGTCTTGGGTT
Db	275	TTCCATTTTAC	CTTGTAG	TTCTTGG	TGATG	CTTGT	GATG	AACTT	GCAGTCTTGGGTT
QY	301	CTGATGTGCT	TTTGG	CCCATCT	GCATG	CTCTG	TTTGT	GATG	ATGCAACATGCTTG
Db	335	CTGATGTGCT	TTTGG	CCCATCT	GCATG	CTCTG	TTTGT	GATG	ATGCAACATGCTTG
QY	361	GACAGGGTAG	TTTCAAG	TGGTGG	TGCTG	CTG	CGGTT	TAC	GACGCTGGCA
Db	395	GACAGGGTAG	TTTCAAG	TGGTGG	TGCTG	CTG	CGGTT	TAC	GACGCTGGCA
QY	421	TTTGTCAAG	CCCTGC	ATCA	CAACAT	CTCT	CTG	ATG	ACCTGGAGTCTTGGCACTGCA
Db	455	TTTGTCAAG	CCCTGC	ATCA	CAACAT	CTCT	CTG	ATG	ACCTGGAGTCTTGGCACTGCA
QY	481	CTGCTCAT	CGCAG	CTTAA	GAGG	TGTG	CAACAT	CGT	CTTAAAGCTGGGCTCTTTC

```
Db 515 CTGCTTGTTCAGAGAGCTGAAGAGGTGTGACAAATGTGCTGTGTGTTTAAAGTCGGGCTCTTC 574
Qy 541 TCGGGCTCTGTGTGACCTGGCCCTGTCTGTCTGTGATCACTGACCGAGCTTTCTGGGAG 600
Db 575 TCTGGCTCTGTGTGATCTGTGCTCTCTCTGTCTGTGATCACTGACCGAGCTTTCTGTGAG 634
Qy 601 CTGCTGTCTGTCACTTCACTTCCCTTACCTGTGCTGATGTGCTGATGTGCTGATGTGCTGCTT 660
Db 635 CTGCTCTCTCTTCTTCACTTCCCTTACCTGTGCTGATGTGCTGATGTGCTGATGTGCTGCTT 694
Qy 661 GCTGCTTACCTGGGCTGTGATGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 720
Db 695 GCTTGTGATCTGGGCTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 754
Qy 721 CAAGGCTCTGTCACTCAAGTCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 780
Db 755 CAAGTCTGATCATCATGATCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 814
Qy 781 GTGTCCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
Db 815 GTGTCCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862

RESULT 6
BC059819
LOCUS
DEFINITION
Mus musculus cancer related gene-liver 1, mRNA linear ROD 21-OCT-2003
MGC:69583 IMAGE:6839525, complete cds.
ACCESSION
BC059819
VERSION
BC059819.1 GI:37590519
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4078)
Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carminci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
McKernan, R.D., Mullahy, S.J., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinaki, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 4078)
Srausberg, R.
Direct Submission
Submitted (07-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNI)
DNA Sequencing by: Genome Sequence Centre,
```

```
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Aeano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Nataaja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 132 Row: 0 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21314857.

FEATURES
            location/Qualifiers
            1..4078
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="MGC:69583 IMAGE:6839525"
            /tissue_type="Brain"
            /clone_lib="NIH_BMAP_GH0"
            /lab_host="DH10B"
            /note="Vector: pYX-ASC"
            1..4078
            /gene="CRG-L1"
            /db_xref="LocusID:230379"
            76..765
            /gene="CRG-L1"
            /codon_start=1
            /product="CRG-L1 protein"
            /protein_id="AAH59819.1"
            /db_xref="GI:37590520"
            /db_xref="LocusID:230379"
            /translation="MGAPWHDHLAGSSEVDWEDNYTIVPAIAEFTNTISNVLFET
            LPICMCLFQYATCNFSGIYLITLVLVVVIGSVYFHTLSFLGOMLDELAILWLM
            CALAMPFPRYLKPIFENDRCNVRVFKLGLFSLWTLALFCWISDPAFCLELSSFH
            FPIHLCHVWHILICLASVLCVCYFAYFDAEIEPQGPVIRFPSEKFAFVGFVSVLL
            CAHKSKSPVKIT"

ORIGIN
Query Match          50.4%;   Score 417.2;   DB 10;   Length 4078;
Best Local Similarity 73.9%;   Pred. No. 2.8e-102;
Matches 612;   Conservative 0;   Mismatches 78;   Indels 138;   Gaps 1;

Qy 1  ATGGGGCCCGCCGCACTGGTGGGACCACTGCAGGCTGGTAGCTCGGAGGTGGACTGGTGC 60
Db 76  ATGGGGCCCGCCGCACTGGTGGGACCACTGCAGGCTGGTAGCTCGGAGGTGGACTGGTGC 135
Qy 61  GAGGAGCACTACACCATCGTGCCTGCTATGCCGAGTCTTACACAGATCAGCAATGTC 120
Db 136  GAGGAGCACTACACCATCGTGCCTGCTATGCCGAGTCTTACACAGATCAGCAATGTC 195
Qy 121  TTATTTTTCATTTTACCGCCCATCTGCATGTGCTTGTGTTGATGAGTATGCAACATGCTTG 180
Db 196  TTGTTTTCATTTTACCTCCCACTGCATGTGTTGTTGCGCCAGTACGACAGTGTTC 255
Qy 181  AACAGTGACATCTACTTAATCTGACATCTTTTGGTGTCTAGTGGGAATGGATCCGTCTAC 240
Db 256  AACAGCGCATCTACTTAATCTGACATCTTCTAGTTCGAGGATGGATCTGTCTAC 315
Qy 241  TTCCATTTTACCTTAGTTTCTTGGGTGCAGATGCTTGCATGATGATGCTTGTGCTTGGGT 300
Db 316  TTCCATGCAACGCTGAGTTTCTTGGGTGCAGATGCTTGCATGATGATGCTTGTGCTTGGGT 375
Qy 301  CTGATGTGTGCTTTGGCCATGTGGTTCCTCCAGAGGTATCTTACCAAGATCTTTTCGGAAT 360
Db 376  CTGATGTGTGCTTTGGCCATGTGGTTCCTCCAGAGGTATTTACCAAGATCTTTTCGGAAT 435
Qy 361  GACAGGGGTAGTTCAAGTGGTGGTGCAGTGTCTCTGTCTCGGCTTACAGCGTGCCTGCA 420
```


KEYWORDS	Homo sapiens (human)	CDS	/gene="bwa"
SOURCE	Homo sapiens		195..1046
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		/gene="bwa"
REFERENCE	1		/notes="ceramide; BWA"
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		/codon_start=1
TITLE	Kits, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof		/product="brain washing"
JOURNAL	Patent: WO 02068579-A 22840 06-SEP-2002;		/protein_id="AAL55991.1"
PE CORPORATION (NY) (US)	PE Corporation		/db_xref="GI:18028135"
FEATURES	Location/Qualifiers		/translation="MGMGGGGLLDIYAMAEQIRPGSSPYDWCENYLISSNIAEFV
source	1..315		NTPSNFLPILLPVLLIMLFKEYGRFVPGHIVLWLVVGLSSMYHATLSLIGQLL
	/organism="Homo sapiens"		DELAIIWFMFAEFLPKRYKPKFVNDKRTESWMLLSAIAATGSLWKKPIVNAEV
	/mol_type="unassigned DNA"		LMFMSVPTVMVLYTELQVRSDQRYRLGRSTTVMAVAVFCVNDIRLFCFAMSSINFP
	/db_xref="taxon:9606"		YLHGFWHIFIFIAAYTVLVLFAVYVESELPPQRLPKLYWPKNEFERGIPFISIRNPG
ORIGIN			KALNTI"
	Query Match 37.3%; Score 308.6; DB 6; Length 315;		
	Best Local Similarity 98.7%; Pred. No. 8.4e-73;		
	Matches 311; Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
QY	514 ATGCGTGTGTTAAGCTGGGCTCTTCTCGGGCTCTGTGGACCTGGCCCTGTTCTGC 573		
Db	1 ATGCGTGTGTTAAGCTGGGCTCTTCTCGGGCTCTGTGGACCTGGCCCTGTTCTGC 60		
QY	574 TGGATCAGTACCGAGCTTTCTGGAGCTGCTGTCATCTTCAACTTCCCTACCTGCAC 633		
Db	61 TGGATCAGTACCGAGCTTTCTGGAGCTGCTGTCATCTTCAACTTCCCTACCTGCAC 120		
QY	634 TGCATGCGGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 693		
Db	121 TGCATGCGGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180		
QY	694 TTTGATGCTGCTCAGAGATCTCTGAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 753		
Db	181 TTTGATGCTGCTCAGAGATCTCTGAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 240		
QY	754 AATGGGCTTCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813		
Db	241 AATGGGCTTCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300		
QY	814 GTCAAGATCAGTGA 828		
Db	301 GTCAAGATCAGTGA 315		
RESULT 10			
AF323976			
LOCUS	AF323976 1163 bp mRNA linear INV 01-JAN-2002		
DEFINITION	Drosophila melanogaster brain washing (bwa) mRNA, complete cds.		
ACCESSION	AF323976		
VERSION	AF323976.1 GI:18028134		
KEYWORDS			
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
	Ephydroidea; Drosophilidae; Drosophila.		
	1 (bases 1 to 1163)		
	Pascual, A., Boquet, I. and Preat, T.		
	Direct Submission		
REFERENCE	Submitted (24-NOV-2000) Institut de Neurobiologie Alfred Fessard,		
AUTHORS	Avenue de la Terrasse, Bat 32/33, Gif-sur-Yvette 91190, France		
TITLE			
JOURNAL			
FEATURES	Location/Qualifiers		
source	1..1163		
	/organism="Drosophila melanogaster"		
	/mol_type="mRNA"		
	/db_xref="taxon:7227"		
	/chromosome="2"		
	/map="38A"		
	1..1163		
gene			

```

RESULT 11
AY071232      2318 bp mRNA linear INV 20-DEC-2001
Drosophila melanogaster RE26924 full length cDNA.
ACCESSION
AY071232
AY071232.1 GI:17945605
FLI CDNA.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 2318)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
Champe, R., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,
George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G.,
Miranda, A., Mungall, C. J., Nunoo, J., Pacleab, J., Paragas, V., Park, S.,
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.
and Celniker, S.
Direct Submission
Submitted (18-DEC-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
FEATURES
source
1..2318
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="38B4-38B5"
1..2318
/gene="CG13969"
/note="alignment with genomic scaffold AB003665. gene does
not completely overlap longest ORF"
/db_xref="FLYBASE:FBgn0032842"
404..1255
/gene="CG13969"
/note="Longest ORF"
/codon_start=1
/product="RE26924p"
/protein_id="AAL48854.1"
/db_xref="GI:17945606"
/db_xref="FLYBASE:FBgn0032842"
/translation="MGNGGGGLDIYAMANEHLRPGSSPVDWCEGNVLISSNTAEFV
NFSNFDLILPPVLMIFKEYGRFVPIGHIWVLLVGLSSMYFPHATLSLIGQLL
DELAIVFMFAALFLFYKRYPKFNDKRTFVWMLLSAIAATGLSKWPIVNAFV
LMMFSVPTWMLYTELQDSQVRVLGIRSTVMAVAVFCWINDRIFCEAWSNFNP
YLHGFWHIFIAATVVLVLFAYFVSELSRQLPQLKWKPNFEFGIPISIRNPG
KALRNTI"
ORIGIN
Query Match 28.8%; Score 238.2; DB 3; Length 2318;
Best Local Similarity 57.4%; Pred. No. 1.5e-53;
Matches 429; Conservative 0; Mismatches 318; Indels 0; Gaps 0;

```


REFERENCE 2 (bases 1 to 1429)
AUTHORS Satou, Y. and Satoh, N.
TITLE Direct Submission
JOURNAL Zoology; Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail: satoh@acidian.zool.kyoto-u.ac.jp, Tel: 81-75-753-4095,
Fax: 81-75-705-1113)
COMMENT Clona intestinalis cDNA Project (URL:
http://ghost.zool.kyoto-u.ac.jp/indexr1.html).
FEATURES
source
1. 1429
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb013nl4"
ORIGIN
Query Match 26.2%; Score 216.6; DB 3; Length 1429;
Best Local Similarity 56.2%; Pred. No. 1.1e-47;
Matches 433; Conservative 0; Mismatches 329; Indels 9; Gaps 1;
Qy 39 TAGCTCGAGGTGGACTGGTGGAGGACAACTACACCATCGTGGCTCTATGCGCGAGTT 98
Db 115 TAGTTCTGAAGTGGACTGGTGGAGAACAAATACGCAACAATACCAATATTGCGGAATT 174
Qy 99 CTACACACATCAGCAATGCTTTATTTTCATTTTACCGGCCATCGCATGCTGTT 158
Db 175 CTGGAACATGTAAAGCAATTCATATTTTATCATTTCCACCCCTGCTACCTGTT 234
Qy 159 TGATGAGTATGCAACATGCTTGAACAGTGCTATCTTAATCTGGACTCTTTGGTGT 218
Db 235 CAAGCAATATCTCGACAAGTGCTCATCAGTTAACTGTGTGGGTTTACTCATATT 294
Qy 219 AGTGGGAATTTGATCCGCTCTACTTCCATTTTACCCTTAGTTTCTGGTGCAGATGCTTGA 278
Db 295 TGTGTGGTGGCTCGCTGATTTTCTCACTCAACTTAAGCCTTGTGGCAGTTAAATGA 354
Qy 279 TGAATTCAGTCTCTTTGGGTTCTGATGTGTGCTTTGGCCATGTTTCCCGAAGGTA 338
Db 355 TGAATATGCAATACTATGGGTGTCTCGCTGCAATTAGCCACTTGGCTACCAAGCAATA 414
Qy 339 TCTACCAAGATCTTTCCGAATGACAGGGTAGGTTCAGGTGGTGGTGCAGTGCCTGTC 398
Db 415 TCTCCCTTCGATCTCAGATCAGACAGCGGAAATTCAGACTAATATAGCTGGTGTGC 474
Qy 399 TGGCGTTACGAGTGGCTGGCATTTGTCAAGCCTGCCATCAACAACATCTCTCTGATGAC 458
Db 475 TGTAGTCACTACTTGTCTTGTCTGTGAACCGGAACTTAACACCGACTATTGTTCAT 534
Qy 459 CTTGGGAGTTCCTTGCATCGACTGCTCTCATCGAGAGCTAAGAGGTGACACATGCG 518
Db 535 GTTTGGGGTACCGAGAACTGATTTGGTGTGTAGAGTTGAGAAGTTGCGAGTCCCGTT 594
Qy 519 TGTGTTTAAAGCTGGGCTCTTCTCGGGCTCTGCTGGACCTGGCCCTGTTCTCTGGAT 578
Db 595 GTTCTACATGTCGGTCAATGTCAGCTATATGTTGGTGGTGGTGGTGGTGGTGGT 654
Qy 579 CAGTGACCGAGCTTTCTGGAGCTGCTGTCATCTTCACTTCCCTTACCTGCACTGAT 638
Db 655 CAGTGATAGATTTTGTGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
Qy 639 GTGGCACAATCTCATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Db 706 ATGGCATAATCAAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
Qy 699 TGTGCTCTCAGAGATTCCTGAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
Db 766 CGCAACAAACGAGTCCCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 825
Qy 759 GGCCTTCATTTGGTTCCTTATGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809
Db 826 TACATGGTTCGGCATTCGGTACGTTTCCCTGTTAGTCACAAAGCGTCAAGTC 876

RESULT 13
AL158206
LOCUS

DEFINITION Human DNA sequence from clone RP11-363E7 on chromosome 9, complete sequence.

ACCESSION AL158206
VERSION AL158206.8 GI:8977646
KEYWORDS HTG.

SOURCE

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 163542)
Laird, G.

REFERENCE

AUTHORS Direct Submission
TITLE Submitted (25-JUN-2000) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

COMMENT

On Jul 7, 2000 this sequence version replaced gi:8651879.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9RP11-363E7 is from the library RPCI-11.2 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pBACe3.6IMPORTANT: This sequence is not the entire insert of clone
RP11-363E7 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.The true left end of clone RP11-363E7 is at 1 in this sequence. The
true left end of clone RP11-25202 is at 163443 in this sequence.

FEATURES

source

Location/Qualifiers
1. 163542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-363E7"
/clone_lib="RP11-11.2"
127756..127796
/note="match: GSS: Em:AG026881"

misc_feature

ORIGIN

Query Match 22.7%; Score 188; DB 9; Length 163542;
Best Local Similarity 100.0%; Pred. No. 9.8e-40;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 641 GGACATCCTCATCTGCTTGTCTGCTTACCTGGCTGTGTATGCTTTGCTTACTTGTATG 700

Db 17329 GGACATCCTCATCTGCTTGTCTGCTTACCTGGCTGTGTATGCTTTGCTTACTTGTATG 17388

Qy 701 CTGCTCAGAGATTCCTGAGCAAGGCCCTGTCAATCAAGTTCTGGCCCAATGAGAAATGG 760

Db 17389 CTGCTCAGAGATTCCTGAGCAAGGCCCTGTCAATCAAGTTCTGGCCCAATGAGAAATGG 17448

	Matches	412;	Conservative	0;	Mismatches	359;	Indels	3;	Gaps	1;
QY	40	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA
Db	25	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA	AGCTCGGAGTGA
QY	100	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC
Db	85	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC	TACAAACAGATC
QY	160	GATGAGTATGCA	GATGAGTATGCA	GATGAGTATGCA	GATGAGTATGCA	GATGAGTATGCA	GATGAGTATGCA	GATGAGTATGCA	GATGAGTATGCA	GATGAGTATGCA
Db	145	CACCCGTATGCC	CACCCGTATGCC	CACCCGTATGCC	CACCCGTATGCC	CACCCGTATGCC	CACCCGTATGCC	CACCCGTATGCC	CACCCGTATGCC	CACCCGTATGCC
QY	220	GTGGGAATGGAT	GTGGGAATGGAT	GTGGGAATGGAT	GTGGGAATGGAT	GTGGGAATGGAT	GTGGGAATGGAT	GTGGGAATGGAT	GTGGGAATGGAT	GTGGGAATGGAT
Db	205	ATAGGCGTGTTC	ATAGGCGTGTTC	ATAGGCGTGTTC	ATAGGCGTGTTC	ATAGGCGTGTTC	ATAGGCGTGTTC	ATAGGCGTGTTC	ATAGGCGTGTTC	ATAGGCGTGTTC
QY	280	GAACCTGCACTC	GAACCTGCACTC	GAACCTGCACTC	GAACCTGCACTC	GAACCTGCACTC	GAACCTGCACTC	GAACCTGCACTC	GAACCTGCACTC	GAACCTGCACTC
Db	265	GAGATGCCATCT	GAGATGCCATCT	GAGATGCCATCT	GAGATGCCATCT	GAGATGCCATCT	GAGATGCCATCT	GAGATGCCATCT	GAGATGCCATCT	GAGATGCCATCT
QY	340	CTACCAAGATCT	CTACCAAGATCT	CTACCAAGATCT	CTACCAAGATCT	CTACCAAGATCT	CTACCAAGATCT	CTACCAAGATCT	CTACCAAGATCT	CTACCAAGATCT
Db	325	TTCCCTCTCTCT	TTCCCTCTCTCT	TTCCCTCTCTCT	TTCCCTCTCTCT	TTCCCTCTCTCT	TTCCCTCTCTCT	TTCCCTCTCTCT	TTCCCTCTCTCT	TTCCCTCTCTCT
QY	400	GCGGTTACGAGT	GCGGTTACGAGT	GCGGTTACGAGT	GCGGTTACGAGT	GCGGTTACGAGT	GCGGTTACGAGT	GCGGTTACGAGT	GCGGTTACGAGT	GCGGTTACGAGT
Db	385	GTGGTCAGCACC	GTGGTCAGCACC	GTGGTCAGCACC	GTGGTCAGCACC	GTGGTCAGCACC	GTGGTCAGCACC	GTGGTCAGCACC	GTGGTCAGCACC	GTGGTCAGCACC
QY	460	CTGGGAGTTCCT	CTGGGAGTTCCT	CTGGGAGTTCCT	CTGGGAGTTCCT	CTGGGAGTTCCT	CTGGGAGTTCCT	CTGGGAGTTCCT	CTGGGAGTTCCT	CTGGGAGTTCCT
Db	445	ATTGCCCTGCAC	ATTGCCCTGCAC	ATTGCCCTGCAC	ATTGCCCTGCAC	ATTGCCCTGCAC	ATTGCCCTGCAC	ATTGCCCTGCAC	ATTGCCCTGCAC	ATTGCCCTGCAC
QY	520	GTGTTTAAGCTG	GTGTTTAAGCTG	GTGTTTAAGCTG	GTGTTTAAGCTG	GTGTTTAAGCTG	GTGTTTAAGCTG	GTGTTTAAGCTG	GTGTTTAAGCTG	GTGTTTAAGCTG
Db	505	CTTCGGCACCCT	CTTCGGCACCCT	CTTCGGCACCCT	CTTCGGCACCCT	CTTCGGCACCCT	CTTCGGCACCCT	CTTCGGCACCCT	CTTCGGCACCCT	CTTCGGCACCCT
QY	580	AGTGACCGAGCT	AGTGACCGAGCT	AGTGACCGAGCT	AGTGACCGAGCT	AGTGACCGAGCT	AGTGACCGAGCT	AGTGACCGAGCT	AGTGACCGAGCT	AGTGACCGAGCT
Db	565	AGTGACCGCTGT	AGTGACCGCTGT	AGTGACCGCTGT	AGTGACCGCTGT	AGTGACCGCTGT	AGTGACCGCTGT	AGTGACCGCTGT	AGTGACCGCTGT	AGTGACCGCTGT
QY	640	TGGCACAATCCT	TGGCACAATCCT	TGGCACAATCCT	TGGCACAATCCT	TGGCACAATCCT	TGGCACAATCCT	TGGCACAATCCT	TGGCACAATCCT	TGGCACAATCCT
Db	625	TGGCATGTGTCT	TGGCATGTGTCT	TGGCATGTGTCT	TGGCATGTGTCT	TGGCATGTGTCT	TGGCATGTGTCT	TGGCATGTGTCT	TGGCATGTGTCT	TGGCATGTGTCT
QY	700	GCTGCTTCAGAT	GCTGCTTCAGAT	GCTGCTTCAGAT	GCTGCTTCAGAT	GCTGCTTCAGAT	GCTGCTTCAGAT	GCTGCTTCAGAT	GCTGCTTCAGAT	GCTGCTTCAGAT
Db	685	GCCAACTATGAT	GCCAACTATGAT	GCCAACTATGAT	GCCAACTATGAT	GCCAACTATGAT	GCCAACTATGAT	GCCAACTATGAT	GCCAACTATGAT	GCCAACTATGAT
QY	760	GCCTTCATTGCT	GCCTTCATTGCT	GCCTTCATTGCT	GCCTTCATTGCT	GCCTTCATTGCT	GCCTTCATTGCT	GCCTTCATTGCT	GCCTTCATTGCT	GCCTTCATTGCT
Db	745	CC---CGTGGG	CC---CGTGGG	CC---CGTGGG	CC---CGTGGG	CC---CGTGGG	CC---CGTGGG	CC---CGTGGG	CC---CGTGGG	CC---CGTGGG

Search completed: November 21, 2004, 03:07:27
Job time : 3363.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:48:16 ; Search time 351.193 Seconds
(without alignments)
12376.463 Million cell updates/sec

Title: US-10-017-410-3

Perfect score: 828

Sequence: 1 atggggcccccgcactgggtg.....catcagtcagatcacgtga 828

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 23Sep04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	100.0	828	6 ABS54149	ABS54149 Human cDN
2	813.6	98.3	4202	6 ACC50961	ACC50961 Human bla
3	813.6	98.3	4212	11 ADN38833	Adn38833 Cancer/an
4	761	91.9	822	4 AHA48641	Aha48641 Human cer
5	664.8	80.3	4175	6 ABS54148	ABS54148 Mouse cDN
6	486.2	58.7	1623	10 ADI21948	Adi21948 Novel hum
7	414.8	50.1	1527	10 ADL06640	Adl06640 Human 3T3
8	414.8	50.1	1527	10 ADL06642	Adl06642 Human 3T3
9	313	37.8	487	10 ADI21468	Adi21468 Novel hum
10	185	22.3	1215	4 AKK51566	Akk51566 Human pol
11	185	22.3	1631	4 AAK52550	Aak52550 Human pol
12	183.4	22.1	792	4 AHA48639	Aha48639 Human cer
13	147.8	17.9	537	4 ABL24043	Abi24043 Drosophil
14	95.6	11.5	4071	4 ABL30432	Abi30432 Drosophil
15	75	9.1	4758	4 ABL24042	Abi24042 Drosophil
16	52.2	6.3	35425	10 ADC87616	Adc87616 Human GPC
17	44.8	5.4	2000	8 ADA71938	Ada71938 Rice gene
18	41.2	5.0	2000	8 ADA71938	Ada71938 Rice gene
19	40.2	4.9	354	10 ABX74425	Abx74425 Human cDN
20	40.2	4.9	636	6 ABL90053	Abi90053 Human pol
21	40.2	4.9	748	4 AAH07955	Aah07955 Human cDN

22	40.2	4.9	801	4	AAH48640	Aah48640 Human cer
23	40.2	4.9	1063	6	ABZ11993	Abz11993 Human pol
24	40.2	4.9	1063	12	ADM44511	Adm44511 Novel hum
25	40.2	4.9	2058	4	AAH15072	Aah15072 Human cDN
26	40.2	4.9	2744	3	AAA96499	Aaa96499 cDNA enco
27	40.2	4.9	4646	5	ABV27900	Abv27900 Human pro
28	40.2	4.9	4646	5	ABV28037	Abv28037 Human pro
29	40.2	4.9	4646	5	ABV22202	Abv22202 Human pro
30	40.2	4.9	6114	4	ABL24565	Abi24565 Drosophil
31	40.2	4.9	9287	4	ABL24564	Abi24564 Drosophil
32	38.8	4.7	332	12	ACH93876	Ach93876 Human gen
33	38.8	4.7	579	12	ACH80176	Ach80176 Human gen
34	38.8	4.7	1057	8	ACA57321	Aca57321 Human adi
35	38.8	4.7	2271	4	AAH14050	Aah14050 Human cDN
36	38.8	4.7	2591	6	ABA93743	Abag93743 Human tra
37	38.8	4.7	3383	6	ABQ60889	Abq60889 FLJ20756
38	38.8	4.7	34989	10	ADC86916	Adc86916 Human GPC
39	38.6	4.7	639	5	AA579600	Aas79600 DNA enco
40	38.6	4.7	10732	3	AAA10594	Aaa10594 Gene enco
41	38	4.6	2800	4	AAAD07131	Aad07131 Canine re
42	38	4.6	2803	4	AAAD07132	Aad07132 Canine re
43	38	4.6	2805	4	AAAD07130	Aad07130 Canine re
44	37.4	4.5	1551	3	AAZ98176	Aaz98176 Human sig

ALIGNMENTS

RESULT 1
ABS54149
ID ABS54149 standard; cDNA; 828 BP.
XX
AC ABS54149;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human cDNA encoding a liver tumour marker protein, CRG-L1.
XX
KW Human; ss; gene; liver cancer; liver tumour; CRG-L1;
KW hepatocellular cancer; chromosome 9p.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..828
FT /*tag= a
FT /product= "Liver tumour marker protein, CRG-L1"
XX
PN US2002115094-A1.
XX
PD 22-AUG-2002.
XX
PF 14-DEC-2001; 2001US-00017410.
XX
PR 14-DEC-2000; 2000US-0255674P.
XX
PA (FARN/) FARNHAM P J.
PA (GRAV/) GRAVEEL C R.
XX
PI Farnham PJ, Graveel CR;
XX
DR WPI; 2002-706409/76.
DR P-PSDB; ABG32881.
XX
PT Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
PT liver cancer, is differentially expressed in liver tumors relative to
PT normal liver tissues.
XX
PS Claim 2; Page 8; 11pp; English.
XX
CC The invention relates to a polypeptide designated as CRG-L1, which is
CC differentially expressed in liver tumours relative to normal expression

Db 327 TTCATGCAACCCCTAGTTCTTGGGTGTCAGATGCTTGATGAACCTGTCAGTCCCTTTGGGTT 386
QY 301 CTGATGTGCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTTCGGAAT 360
Db 387 CTGATGTGCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTTCGGAAT 446
QY 361 GACAGGGTAGGTTCAAGGTGGTGGTCAAGTGTCTGTCGGGTACGAGTGCCTGGCA 420
Db 447 GACCGGGTAGGTTCAAGGTGGTGGTCAAGTGTCTGTCGGGTACGAGTGCCTGGCA 506
QY 421 TTTGTCAAGCTGTCATCAACCAACATCTCTGATGACCCCTGGGAGTTCCTTGCACTGCA 480
Db 507 TTTGTCAAGCTGTCATCAACCAACATCTCTGATGACCCCTGGGAGTTCCTTGCACTGCA 566
QY 481 CTGCTATCGCAGAGCTAAAGAGGTGGAACAATGCGTGTGTTTAAAGCTGGGCTCTTC 540
Db 567 CTGCTATCGCAGAGCTAAAGAGGTGGAACAATGCGTGTGTTTAAAGCTGGGCTCTTC 626
QY 541 TCGGGCTCTGGTGGACCCCTGGCCCTGTTCTGCTGGATCAGTGACCGAGCTTTCTGCGAG 600
Db 627 TCGGGCTCTGGTGGACCCCTGGCCCTGTTCTGCTGGATCAGTGACCGAGCTTTCTGCGAG 686
QY 601 CTGCTGTATCTCTCAACTTCCCTACCTGCACTGCAATGCGGACATCTCTCATCTGCTT 660
Db 687 CTGCTGTATCTCTCAACTTCCCTACCTGCACTGCAATGCGGACATCTCTCATCTGCTT 746
QY 661 GCTGCTTACCTGGGCTGTATGCTTTGCTTACCTTGTATGCTGCTCAGAGATTCCTGAG 720
Db 747 GCTGCTTACCTGGGCTGTATGCTTTGCTTACCTTGTATGCTGCTCAGAGATTCCTGAG 806
QY 721 CAAGGCCCTGTCTATCAAGTTCGCGCCCAATGAGAAATGAGAAATGGGCTTCATTGGTGCTCCCTAT 780
Db 807 CAAGGCCCTGTCTATCAAGTTCGCGCCCAATGAGAAATGAGAAATGGGCTTCATTGGTGCTCCCTAT 866
QY 781 GTGTCCCTCTGTGTGCGCAACAAGAAATCATCAGTCAAGATCAGGTGA 828
Db 867 GTGTCCCTCTGTGTGCGCAACAAGAAATCATCAGTCAAGATCAGGTGA 914

RESULT 4

AAH48641

ID AAH48641 standard; DNA; 822 BP.

AC AAH48641;

DT 21-SEP-2001 (first entry)

DE Human ceramidase K3 DNA.

KW Ceramidase; human; K3; antiproliferative; anticancer; anti-eczema;
KW antipsoriasis; dermatological; ceramide; sphingosine; treatment; ds;
KW altered cell proliferation; cancer; neurodermatitis; eczema; cosmetic;
KW permeability barrier; psoriasis; ichthyosis; lamellar ichthyosis ICR2.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 1..822

FT /*tag= a

FT /transl_except= (pos:382..387,aa:Cys)

FT /note= "No stop codon given"

XX WQ200155408-A1.

XX 02-AUG-2001.

XX 27-JAN-2001; 2001WO-EP000900.

XX 27-JAN-2000; 2000DE-01003293.

XX 09-MAR-2000; 2000DE-01011392.

XX (MEMO-) MEMOREC MEDICAL MOLECULAR RES COLOGNE ST.

XX

PI Hofmann K, Conradt M;
XX WPI; 2001-483256/52.
DR P-PSDB; AAB86365.
XX New ceramidase containing specific structural motif, useful for
PT diagnosis, prevention and treatment of ceramidase defects, e.g.
PT ichthyosis, also in cosmetics.
XX
PS Claim 3; Page 17; 31pp; German.
XX This invention describes novel human ceramidase (I) containing a specific
CC structural motif which has antiproliferative, anticancer, anti-eczema,
CC antipsoriasis and dermatological activity. (I) cleave ceramide, resulting
CC in formation of sphingosine. (II), or the nucleic acid (II) encoding it,
CC are used for diagnosis, prevention or treatment of diseases associated
CC with ceramide defects, particularly altered cell proliferation (cancer)
CC or altered ceramide layers on the skin (neurodermatitis, eczema,
CC psoriasis), also for targeted modification of the permeability barrier by
CC ceramidase or its activators, e.g. for transcutaneous delivery of
CC substances. Specifically they are used for diagnosis of ichthyosis,
CC particularly lamellar ichthyosis ICR2. (I) can also be used in cosmetics.
XX This sequence encodes human ceramidase K3
SQ Sequence 822 BP; 153 A; 225 C; 209 G; 235 T; 0 U; 0 Other;

Query Match

Best Local Similarity 91.9%; Score 761; DB 4; Length 822;

Matches 792; Conservative 0; Mismatches 30; Indels 3; Gaps 1;

QY 1 ATGGGGCCCGCAGCTGGTGGGACCGAGCTGCAGCTGGTAGCTGGAGGTGGACTGCTGC 60
Db 1 ATGGGGCCCGCAGCTGGTGGGACCGAGCTGCAGCTGGAGGTGGACTGCTGC 60
QY 61 GAGGACAATACACCATCTGTCGCTGCTATCGCGAGTCTTCAACACGATCAGCAATGTC 120
Db 61 GAGGACAATACACCATCTGTCGCTGCTATCGCGAGTCTTCAACATGATCAGCAATGTC 120
QY 121 TTAATTTTCAATTTACCGCCCATCTGATGCTGCTTTGTTGATGAGTATGCAATGCTTG 180
Db 121 TTAATTTTCAATTTACCGCCCATCTGATGCTGCTTTGTTGATGAGTATGCAATGCTTG 180
QY 181 AACAGTGACATCTACTTAATCTGACCTCTTTGTTGTTAGTGGGAAATGGATCCGTCTAC 240
Db 181 AACAGTGACATCTACTTAATCTGACCTCTTTGTTGTTAGTGGGAAATGGATCCGTCTAC 237
QY 241 TTCCATTTTACCCTTAGTCTTTTGGGTTCAGATGCTTGAATGAACTTTCAGTCTTTGGGTT 300
Db 238 TTCCATGCAACCCCTTAGTCTTTTGGGTTCAGATGCTTGAATGAACTTTCAGTCTTTGGGTT 297
QY 301 CTGATGTGCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTTCGGAAT 360
Db 298 CTGATGTGCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTTCGGAAT 357
QY 361 GACAGGGTAGGTTCAAGGTGGTGGTCAAGTGTCTGTCGGTTACGAGTGCCTGGCA 420
Db 358 GACAGGGTAGGTTCAAGGTGGTGGTCAAGTGTCTGTCGGTTACGAGTGCCTGGCA 417
QY 421 TTTGTCAAGCTGCGCATCAACAAATCTCTGATGACCCCTGGGAGTTCCTTGCACTGCA 480
Db 418 TTTGTCAAGCTGCGCATCAACAAATCTCTGATGACCCCTGGGAGTTCCTTGCGCTGCA 477
QY 481 CTGCTCATCGCAGAGCTAAAGAGGTGGAACAATGCTGTTTAAAGCTGGGCTCTTC 540
Db 478 CTGCTCATCGCAGAGCTAAAGAGGTGGAACAATGCTGTTTAAAGCTGGGCTCTTC 537
QY 541 TCGGGCTCTGGTGGACCCCTGGCCCTGTTCTGCTGGATCAGTGACCGAGCTTTCTGCGAG 600
Db 538 TCGGGCTCTGGTGGACCCCTGGCCCTGTTCTGCTGGATCAGTGACCGAGCTTTCTGCGAG 597
QY 601 CTGCTGTATCTCTCAACTTCCCTACCTGCACTGCAATGCGGACATCTCTCATCTGCTT 660
Db 598 CTGCTGTATCTCTCAACTTCCCTACCTGCACTGCAATGCGGACATCTCTCATCTGCTT 657

QY 661 GCTGCTACCTGGGCTGTGTATGCTTTGGCTTACCTTGTGATGCTGCTCAGAGATTCCTGAG 720
Db |||||
QY 658 GCTGCTACCTGGGCTGTGTATGCTTTGGCTTACCTTGTGATGCTGCTCAGAGATTCCTGAG 717
Db |||||
QY 721 CAAGGCCCTGTCAATCAAGTCTGGCCCAATGAGAAATGGGGCTTCAATGGTGTCCCTAT 780
Db |||||
QY 718 CAAGGCCCTGTCAATCAATCTGGCCCAAGAGAAATGGGGCTTCAATGGTGTCCCTAT 777
Db |||||
QY 781 GTGTCCCTCCTGTGTGCAACAAGAAATCATCAGTCAAGATCAG 825
Db |||||
QY 778 GTGTCCCTCCTGTGTGCAACAAGAAATCATCAGTCAAGATCAG 822
Db |||||

RESULT 5

ABS54148
ID ABS54148 standard; cDNA; 4175 BP.

XX ABS54148;

XX 25-NOV-2002 (first entry)

XX Mouse cDNA encoding a liver tumour marker protein, CRG-L1.

XX Mouse; ss; gene; liver cancer; liver tumour; CRG-L1;
KW hepatocellular cancer.

XX Mus musculus.

XX Key Location/Qualifiers
FT 35..862
FT /*cag= a
FT /product= "Liver tumour marker protein, CRG-L1"

XX US2002115094-A1.

XX 22-AUG-2002.

XX 14-DEC-2001; 2001US-00017410.

XX 14-DEC-2000; 2000US-0255674P.

XX (FARN/) FARNHAM P J.
XX (GRAV/) GRAVEEL C R.

XX Farnham PJ, Graveel CR;

XX WPI; 2002-706409/76.

XX P-PSDB; ABG32880.

XX Novel polypeptide designated as CRG-L1, useful as diagnostic marker for
FT liver cancer, is differentially expressed in liver tumors relative to
FT normal liver tissues.

XX Claim 2; Page 4-7; 11pp; English.

XX The invention relates to a polypeptide designated as CRG-L1), which is
CC differentially expressed in liver tumours relative to normal expression
CC in normal liver tissues, designated CRG-L1. Also included are the
CC encoding polynucleotides (in the case of the human sequence, mapping to
CC chromosome 9p), expression constructs, host cells, anti-CRG-L1
CC antibodies, identifying modulators of CRG-L1, and the use of the CRG-L1
CC sequence in the diagnosis of hepatocellular cancer in tumour cells from a
CC liver of a human or non-human animal. The CRG-L1 protein and
CC polynucleotide are useful as diagnostic markers for a liver cancer in
CC humans and non-human animals, and as a system for assessing putative
CC therapeutic agents. The present sequence encodes mouse CRG-L1

XX Sequence 4175 BP; 962 A; 971 C; 1012 G; 1230 T; 0 U; 0 Other;

XX Query Match 80.3%; Score 664.8; DB 6; Length 4175;

XX Best Local Similarity 87.7%; Pred. No. 9.5e-192;

XX Matches 726; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCACTGGTGGACCACTGGTGGAGTGGTAGCTCGAGGTGGAGTGGTGC 60
Db |||||
QY 35 ATGGGCGCCCGCACTGGTGGACCACTGGTGGAGTGGTAGCTCGAGGTGGAGTGGTGC 94
Db |||||
QY 61 GAGGACAACATACACCATCGTGGCTGCTATCGCCGAGTCTTACACACGATCAGCAATGTC 120
Db |||||
QY 95 GAGGACAACATACACCATCGTGGCTGCTATCGCCGAGTCTTACACACGATCAGCAATGTC 154
Db |||||
QY 121 TTATTTTTCATTTTACCGCCCATCTGCATGTGCTTTGTGATGATGATGCAACATGCTTG 180
Db |||||
QY 155 TTGTTTTTCATTTTACCTCCCATCTGCATGTGCTTTCCGCCAGTACGAACTGCTTC 214
Db |||||
QY 181 AACAGTGACATCTACTTAATCTGGCATCTTTTGGTGTGTAGTGGAAATGGATCCGTCTAC 240
Db |||||
QY 215 AACAGCGGCATCTACTTAATATGACGCTCTAGTGTGTAGTGGGATTTGGATCTGTCTAC 274
Db |||||
QY 241 TTCCATTTTACCTTACTTTTGGGTCAGATGCTTGCATGAACTTGCAGTCTTTTGGGTT 300
Db |||||
QY 275 TTCCATGCAACGCTGAGTTTCTGGGTGAGATGCTTGTGAACTTGGCAATCTGTGGGTT 334
Db |||||
QY 301 CTGATGTGTCTTTGGCCATGTGTTCCCGAAGAGTATCTACCAAGATCTTTTCGGAAT 360
Db |||||
QY 335 CTGATGTGTCTTTGGCCATGTGTTCCCGAAGAGTATTTACCAAGATCTTTTCGGAAT 394
Db |||||
QY 361 GACAGGGGTAGTTCAAGGTGGTGTGCTGTCTGTGCGGTTACGACGTGCTGCGCA 420
Db |||||
QY 395 GACAGGGGCAGGTTCAAGGGCAGTGGTGTGTCTGTCTGCAATTAACAACGCTGTGGCG 454
Db |||||
QY 421 TTTGTGAAGCTGCCATCAACACATCTCTGTATGACCTTGGGAGTCTTTGCACTGCA 480
Db |||||
QY 455 TTTATCAAGCCCGCCATCAACAAATTTTCCCTGATGATTTCTGGGACTTTCCATGCACTGCG 514
Db |||||
QY 481 CTGCTCATCGCAGAGCTAAAGAGGTGTGAACAATGCTGTGTGTTAAAGCTGGGCTCTTC 540
Db |||||
QY 515 CTGCTTGTGCAGAGCTGAAGAGGTGTGAACAATGCTGTGTGTTAAAGCTGGGCTCTTC 574
Db |||||
QY 541 TCGGGCTCTGTGGGACCCCTGGGCTCTTCTGTGCGATCAGTGAACGAGCTTTCTGGGAG 600
Db |||||
QY 575 TCTGGCTCTGTGGGACTCTGGCTCTCTTCTGTGGATCAGCGACCAAGGCTTCTGTGAG 634
Db |||||
QY 601 CTGCTGTCACTCTCAACTTCCCTACCTGCACTGCACTGTGGACATCTCTATCTGCTT 660
Db |||||
QY 635 CTGCTCTCTCTCTTCACTTCCCTACCTGCACTGTGTGGCATATCTCTATCTGCTT 694
Db |||||
QY 661 GCTGCTACCTGGGCTGTGTATGCTTTGGCTTACTTTGATGCTGCTCAGAGATTCCTGAG 720
Db |||||
QY 695 GCTTCTGTAACCTGGGCTGTGTGCTTGGCTTACTTGTGCTGCTCAGAGATCTGAG 754
Db |||||
QY 721 CAAGGCCCTGTCAATCAAGTCTGGCCCAATGAGAAATGGGCTTCAATGGTGTCCCTAT 780
Db |||||
QY 755 CAAGSTCCAGTCAATCAGATTTCTGGCCAGGAGAAATGGGCTTTTATTTGGTGTCCCTAT 814
Db |||||
QY 781 GTGTCCCTCTGTGTGCCCAACAGAAATCATCAGTCAAGATCAGTGA 828
Db |||||
QY 815 GTGTCCCTTCTGTGTGCCCAACAGAAAGTCCGCAAGATCAGTGA 862
Db |||||

RESULT 6

AD121948

ID AD121948 standard; cDNA; 1623 BP.

XX AD121948;

XX 15-APR-2004 (first entry)

XX Novel human protein cDNA #207.

XX forensic; nutritional source; damaged tissue; diseased tissue;

XX myeloid cell disorder; lymphoid cell disorder;

XX bone cartilage tissue growth; tendon tissue growth;

XX ligament tissue growth; nerve tissue growth; regeneration; wound healing;

XX tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;

KW ss; gene.
XX Homo sapiens.
XX WO2003025148-A2.
XX 27-MAR-2003.
XX 19-SEP-2002; 2002WO-US029964.
XX 19-SEP-2001; 2001US-0323739P.
XX 13-SEP-2002; 2002US-00323739.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
XX Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
XX Haley-Vicente D;
XX WPI: 2003-354603/33.
XX P-PSDB; ADI21232.
XX New polynucleotides and secreted proteins, useful for treating myeloid or
XX lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
XX tissue growth or regeneration, in wound healing, and in tissue repair and
XX replacement.
XX Claim 1; SEQ ID NO 207; 156pp; English.
XX The invention relates to an isolated polynucleotide encoding a
XX polypeptide with biological activity. The polynucleotides and
XX polypeptides are useful in diagnostics, forensics, gene mapping,
XX identification of mutations responsible for genetic disorders and other
XX traits, to assess biodiversity, as nutritional sources or supplements.
XX The polynucleotides may also be used as molecular weight markers,
XX chromosome markers or map related gene positions, or as an antigen to
XX raise anti-DNA antibodies or elicit immune response. The polypeptides are
XX useful for raising antibodies, as markers for tissues in which the
XX corresponding polypeptide is expressed, for re-engineering damaged or
XX diseased tissues, for treating myeloid or lymphoid cell disorders, in
XX bone cartilage, tendon, ligament and/or nerve tissue growth or
XX regeneration, in wound healing, in tissue repair and replacement, in
XX healing of burns, incisions and ulcers, and in treating cancer. The
XX present sequence represents cDNA encoding a novel human protein.
XX Sequence 1623 BP; 305 A; 457 C; 382 G; 479 T; 0 U; 0 Other;
SQ
Query Match 58.7%; Score 486.2; DB 10; Length 1623;
Best Local Similarity 97.4%; Pred. No. 2e-137;
Matches 494; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 322 TGGTCCCGAGAGGTATCTACCAAGATCTTCGGAATGACAGGGGTAGGTTCAAGGTG 381
DB 920 TTGTTTCGTCAAAAGGTATCTACCAAGATCTTCGGAATGACAGGGGTAGGTTCAAGGTG 979
QY 382 GTGGTCAGTCTCCTGCTGGGTTTACGAGTGTGCTGCAATTTCTCAAGCTGCGCATCAAC 441
DB 980 GTGGTCTGTCTCTGTCTGAGTTATGACGTGTGCTGCAATTTCTCAAGCTGCGCATCAAC 1039
QY 442 AACATCTCTCTGATGACCCCTGGAGTTCTTTCGACATGCACTGCTCATCGCAGAGCTAAAG 501
DB 1040 AACATCTCTCTGATGACCCCTGGAGTTCTTTCGACATGCACTGCTCATCGCAGAGCTAAAG 1099
QY 502 AGGTGTGACACATGCGTGTGTTTAAAGCTGGGCTCTTCTCGGGCTCTGGTGACCCCTG 561
DB 1100 AGGTGTGACACATGCGTGTGTTTAAAGCTGGGCTCTTCTCGGGCTCTGGTGACCCCTG 1159
QY 562 GCCTGTGTTCTGCTGGATCAGTGACCGAGCTTTCTGGAGCTGTGTGTCATCTTCAACTTC 621
DB 1160 GCCTGTGTTCTGCTGGATCAGTGACCGAGCTTTCTGGAGCTGTGTGTCATCTTCAACTTC 1219
QY 622 CCTTACTGTCAGTGTGCGACATCTCATCTGCGCTTCTGCTGCTACCTGGGCTGTGTA 681

DB 1220 CCTACCTGCATGTCATGTGGCACATCTCTATCTGCTTGCCTACCTGGGCTGTGTA 1279
QY 682 TGTCTTTCCTTACCTTTCATGCTGCTCAGAGATTCCTGAGCAAGGCCCTGTCTCATCAAGTTC 741
DB 1280 TGTCTTTCCTTACCTTTCATGCTGCTCAGAGATTCCTGAGCAAGGCCCTGTCTCATCAAGTTC 1339
QY 742 TGGCCCAATGAGAAATGGGCCCTTCATTTGGTGTGCTCCCTATGTGTGCTCCCTGTGTGCAAC 801
DB 1340 TGGCCCAATGAGAAATGGGCCCTTCATTTGGTGTGCTCCCTATGTGTGCTCCCTGTGTGCAAC 1399
QY 802 AAGAAATCATCAGTCAAGATCAGTGA 828
DB 1400 AAGAAATCATCAGTCAAGATCAGTGA 1426
RESULT 7
ADL06640
ID ADL06640 standard; cDNA; 1527 BP.
XX
AC ADL06640;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human 3T3 cell conversion promoter PP11646 cDNA.
XX
KW 3T3 cell conversion; promoter; human; gene; ss.
XX
OS Homo sapiens.
XX
PN CN1403477-A.
XX
PD 19-MAR-2003.
XX
PF 12-SEP-2001; 2001CN-00126725.
XX
PR 12-SEP-2001; 2001CN-00126725.
XX
PA (SHAN-) SHANGHAI XINSHIJI GENE TECHN DEV CO LTD.
XX
PI Gu J, Yang S;
XX
DR WPI; 2003-494226/47.
DR P-PSDB; ADL06641.
XX
PT Human protein with function of promoting 3T3 cell conversion and its
PT coding sequence.
XX
Example 1; SEQ ID NO 1; 41pp; Chinese.
XX
CC This invention describes a novel human protein with 3T3 cell conversion
CC promoting function, polynucleotides encoding the polypeptide and the
CC recombinant process of producing the polypeptide. The present invention
CC also discloses the agonist resisting the polypeptide and its treatment
CC effect. The present invention also discloses the application of the
CC polynucleotides encoding the human protein with 3T3 cell conversion
CC promoting function.
XX
SQ Sequence 1527 BP; 376 A; 357 C; 383 G; 411 T; 0 U; 0 Other;
Query Match 50.1%; Score 414.8; DB 10; Length 1527;
Best Local Similarity 82.4%; Pred. No. 1.3e-115;
Matches 532; Conservative 0; Mismatches 17; Indels 97; Gaps 1;
QY 108 GATCAGCAATGTCTTATTTTTCATTTTACGCCCATCTGCATGCTGTTTCATGAGTA 167
DB 178 GATCAGCAATGTCTTATTTTTCATTTTACGCCCATCTGCATGCTGTTTCATGAGTA 237
QY 168 TGCAACATGCTTGAACAGATGACATCTACTTAATCTCGACTCTTTTGGTTGTAGTGGGAAT 227
DB 238 TGCAACATGCTTCAACAGATGGCATCTACTTAATCTCGACTCTTTTGGTTGTAGTGGGAAT 297
QY 228 TGATCCGCTACTTCCATTTTACCCCTAGTTTCTTGGGTGAGATGCTTGTGTAACCTTGC 287

298 TGGATCCGCTACTTCCATGCAACCCCTTAGTTTCTTGGTCAGATGCTTGATGAACCTGC 357
288 AGTCCTTTGGGTTCTGATGTCGTCTTGGCCATGTCGTTCCCGAAGGTATCTACCAA 347
358 AGTCCTTTGGGTTCTGATGTCGTCTTGGCCATGTCGTTCCCGAAGGTATCTACCAA 417
348 GATCTTTCCGAATGACAGGGGTAGGTTCAAGGTGTCGTGTCAGTGCCTGCGGTTAC 407
418 GATCTTTCCGAATGACCGGGGTAGGTTCAAGGTGTCGTGTCAGTGCCTGCGGTTAC 477
408 GACGTGCTGTCATTTGTCAGCCCTGCCATCAACATCTCTCTGATGACCCCTGGAGT 467
478 GACGTGCTGTCATTTGTCAGCCCTGCCATCAACATCTCTCTGATGACCCCTGGAGT 537
468 TCCTTGCACTGTCATCTCGCAGAGCTAAAG----- 501
538 TCCTTGCACTGTCATCTCGCAGAGCTAAAGAGGTCATGAGAGGACCGAGGAG 597
502 ----- 501
598 ACACAGGAAGGTGGCCAGCAGGAGGTGGAGACAAGGTCCTGACGATGAGTCTCTG 657
502 -----AGTGTGACAACATGCGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCT 550
658 GACCCCGTGCAGGTGTGACAACATGCGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCT 717
551 GGTGGACCTGGCCCTGTTCTGTCGATCAGTGCAGCGAGCTTCTGGAGTCTGTGTCAT 610
718 GGTGGACCTGGCCCTGTTCTGTCGATCAGTGCAGCGAGCTTCTGGAGTCTGTGTCAT 777
611 CCTTCAACTTCCCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
778 CCTTCAACTTCCCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823

RESULT 8

ADL06642
ID ADL06642 standard; DNA; 1527 BP.
XX
AC ADL06642;
XX
XX 06-MAY-2004 (first entry)
XX Human 3T3 cell conversion promoter Pp11646 DNA.
XX 3T3 cell conversion; promoter; human; gene; ds.
XX Homo sapiens.

Key Location/Qualifiers
CDS 218..640
/*tag= a
/product= "Pp11646"

FN1403477-A.

19-MAR-2003.

12-SEP-2001; 2001CN-00126725.

12-SEP-2001; 2001CN-00126725.

(SHAN-) SHANGHAI XINSHIJI GENE TECHN DEV CO LTD.

Gu J, Yang S;

WPI; 2003-494226/47.

P-PSDB; ADL06641.

Human protein with function of promoting 3T3 cell conversion and its coding sequence.

Example 1; SEQ ID NO 3; 41pp; Chinese.

XX This invention describes a novel human protein with 3T3 cell conversion promoting function, polynucleotides encoding the polypeptide and the recombinant process of producing the polypeptide. The present invention also discloses the agonist resisting the polypeptide and its treatment effect. The present invention also discloses the application of the polynucleotides encoding the human protein with 3T3 cell conversion promoting function.

XX Sequence 1527 BP; 376 A; 357 C; 383 G; 411 T; 0 U; 0 Other;

Query Match 50.1%; Score 414.8; DB 10; Length 1527;
Best Local Similarity 82.4%; Pred. No. 1.3e-115;

Matches 532; Conservative 0; Mismatches 17; Indels 97; Gaps 1;

QY 108 GATCAGCAATGCTCTATTTTTCATTTTACCGCCCATCTGCATGCTGTTGTTGATGACTA 167
DB 178 GATCAGCAATGCTCTATTTTTCATTTTACCGCCCATCTGCATGCTGTTGTTGATGACTA 237
QY 168 TGCACATGCTTGACAGTGCATCTACTTAATCTGCATCTTTTGGTTGTAGTGGGAT 227
DB 238 TGCACATGCTTGACAGTGCATCTACTTAATCTGCATCTTTTGGTTGTAGTGGGAT 297
QY 228 TGGATCCGTCCTACTTCCATTTTACCGCTTACTTCTTGGTTCAGATGCTTGTGATGACTGC 287
DB 298 TGGATCCGTCCTACTTCCATTTTACCGCTTACTTCTTGGTTCAGATGCTTGTGATGACTGC 357
QY 288 AGTCCTTTGGGTTCTGATGTCGTGTTGGCCATGTCGTTCCCGAAGGTATCTACCAA 347
DB 358 AGTCCTTTGGGTTCTGATGTCGTGTTGGCCATGTCGTTCCCGAAGGTATCTACCAA 417
QY 348 GATCTTTCCGAATGACAGGGGTAGGTTCAAGGTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTG 407
DB 418 GATCTTTCCGAATGACCGGGGTAGGTTCAAGGTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTG 477
QY 408 GACGTGCTGTCATTTGTCAGCCCTGCCATCAACATCTCTCTGATGACCCCTGGAGT 467
DB 478 GACGTGCTGTCATTTGTCAGCCCTGCCATCAACATCTCTCTGATGACCCCTGGAGT 537
QY 468 TCCTTGCACTGTCATCTCGCAGAGCTAAAG----- 501
DB 538 TCCTTGCACTGTCATCTCGCAGAGCTAAAGAGGTCATGAGAGGACCGAGGAG 597
QY 502 ----- 501
DB 598 ACACAGGAAGGTGGCCAGCAGGAGGTGGAGACAAGGTCCTGACGATGAGTCTCTG 657
QY 502 -----AGTGTGACAACATGCGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCT 550
DB 658 GACCCCGTGCAGGTGTGACAACATGCGTGTGTTTAAGCTGGGCTCTTCTCGGGCTCT 717
QY 551 GGTGGACCTGGCCCTGTTCTGTCGATCAGTGCAGCGAGCTTCTGGAGTCTGTGTCAT 610
DB 718 GGTGGACCTGGCCCTGTTCTGTCGATCAGTGCAGCGAGCTTCTGGAGTCTGTGTCAT 777
QY 611 CCTTCAACTTCCCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
DB 778 CCTTCAACTTCCCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823

RESULT 9

ADL21468
ID ADL21468 standard; cDNA; 487 BP.

XX
XX ADL21468;

XX AC
XX ADL21468;

DT 15-APR-2004 (first entry)

XX Novel human expressed sequence tag, EST #167.

XX forensic; nutritional source; damaged tissue; diseased tissue;
XX myeloid cell disorder; lymphoid cell disorder;
XX bone cartilage tissue growth; tendon tissue growth;


```
Db 80 AGCTCCGAGGTGGACTGGTGTGAGAGCAACTCCAGTACTCGAGCTGGTGGCCGAGTTC 139
Qy 100 TACAACAGGATCAGCAATGCTTATTTTATTTTACCGCCCAATCTGATGCTGTGTTT 159
Db 140 TACAACAGGATCAGCAATGCTTATTTTATTTTACCGCCCAATCTGATGCTGTGTTT 199
Qy 160 GATGAGTATGCAACATGCTTGAACAGTGCACATCTACTTAACTTGAACATCTTGTGTTA 219
Db 200 CACCGTATGCCAGAGCGCTCCGCTACATTTACGTTGTCTGGGTCTCTTTCATGATC 259
Qy 220 GTGGAAATGGATCCGCTACTCTTCCATTTTACCTTAGTTTCTTGGGTGAGATGCTGTAT 279
Db 260 ATAGGCTGTTCTCAATGATTTTCAACATGAGCTCAGCTTCTTGGGCCAGCTGCTGAC 319
Qy 280 GAACCTTGAGTCTTGTGGTGTCTGATGCTGCTTGGCCATGCTGCTTCCGAGAGGAT 339
Db 320 GAGATCGCATCTCTGTGGCTCTCGGAGTGGCTATAGCATATGATGATCCCGCTGCTAT 379
Qy 340 CTACCAAGATCTTTCGGAATGACAGGGGTAGGTTCAAGGTGGTGGTCAAGTCTCTGTCT 399
Db 380 TTCCTCTCTTCTTGGGGGAAACAGGTTCCAGTTTCATCCGCTGTGTTTCATCACCAC 439
Qy 400 GCGGTTACAGCTGCTGGCATTTGTCAAGCTGCTGCATCAACAACTCTCTCTGATGAC 459
Db 440 GTGGTCAGCACCTTCTGTCTCTCTGCGGCCACGGTCAAGCTTACGCTTCAACAGC 499
Qy 460 CTGGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
Db 500 ATTGCCCTGCACTTCTCTACATGCTGTGCGGAGGTACAGGAAGACCAAGCAATAAGG 559
Qy 520 GTGTTTAACTGGGCTCTTCTGCGGCTCTGCTGCGGCTGCTGCGGCTGCTGCTGCTGCT 579
Db 560 CTGCGGACCTGATGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619
Qy 580 AGTGACCGAGCTTCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639
Db 620 AGTGACCGTCTGCTTTCAGCTTCTGCGAGGATTCATTTCTTCTATCTGACAGCATC 679
Qy 640 TGGCAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699
Db 680 TGGCAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739
Qy 700 GCTGCTCAGAGATCTCTGAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
Db 740 GCAACTATGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799
Qy 760 GCCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
Db 800 CC---CGTGGGCTGCCCTAGTGTGAATCGGGGTGATGACAGGACTGCTGAGAC 853
```

RESULT 11

AAK52550
ID AAK52550 standard; cDNA; 1631 BP.

AC AAK52550;

XX 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 2079.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00683561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX (HYSE-) HYSEQ INC.
Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
P-PSDB; AAM79417.

Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

Claim 1; Page 4462; 6221pp; English.

The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

Sequence 1631 BP; 247 A; 399 C; 344 G; 350 T; 0 U; 291 Other;

Query Match 22.3%; Score 185; DB 4; Length 1631;

Best Local Similarity 53.3%; Pred. No. 2e-45;

Matches 414; Conservative 0; Mismatches 360; Indels 3; Gaps 1;

Qy 40 AGCTCGAGGTGGACTGGTGTGAGAGCAACTCCAGTACTCGAGCTGGTGGCCGAGTTC 99

Db 80 AGCTCGAGGTGGACTGGTGTGAGAGCAACTCCAGTACTCGAGCTGGTGGCCGAGTTC 139

Qy 100 TACAACAGGATCAGCAATGCTTATTTTATTTTACCGCCCAATCTGATGCTGTGTTT 159

Db 140 TACAACAGGATCAGCAATGCTTATTTTATTTTACCGCCCAATCTGATGCTGTGTTT 199

Qy 160 GATGAGTATGCAACATGCTTGAACAGTGCATCTTAACTGCTGACTCTTTTGGTGA 219

Db 200 CACCGTATGCCAGAGCGCTCCGCTACATTTACGTTGTCTGGTCTCTTCAATGATC 259

Qy 220 GTGGAAATGGATCCGCTACTTCCATTTTACCCCTTAGTTTCTTGGGTGAGATGCTTAT 279

Db 260 ATAGGCTGTTCTCCATGATTTTCCACATGAGCTCAGCTTCTTGGGCCAGCTGCTGAC 319

Qy 280 GAACCTTGAGTCTTGTGGTGTCTGATGCTGCTTGTGGCCATGCTGTTCCCGAGAGGAT 339

Db 320 GAGATCGCATCTCTGTGGCTCTCGGAGTGGCTATAGCATATGATGATGCTGCTGCTAT 379

Qy 340 CTACCAAGATCTTTCGGAATGACAGGGGTAGGTTCAAGGTGGTGGTGCAGTGTCTGTCT 399

Db 380 TTCCTCTCTTCTTGGGGGAAACAGGTTCCAGTTTCATCCGCTGGTCTTTCATCACCAC 439

Qy 400 GCGGTTACAGCTGCTGCGCATTTGTCAAGCTGCTGCCATCAACACATCTCTCTGATGAC 459

Db 440 GTGGTCAGCACCTTCTGTCTCTTCTCGGCCCAAGGTTCAAGCTTACGCTTCAACAGC 499

Qy 460 CTGGAGTCTCTTGGCATGCTGCTCATCGAGAGCTAAAGAGGTGTGACAAACATCGGT 519

Db 500 ATTGCCCTGCACATCTCTACATCGTGTGCCAGAGTACAGGACCAACCAATAGGAG 559
 QY 520 GTGTTTAAGCTGGGCTCTTCTCGGGCTCTGGTGACCCCTGGGCCCTGTTCTGCTGATC 579
 Db 560 CTTCCGGCACCTGATGAGGTCTCCGTGGTCTTATGGGCTGTGCTCTGACCACTGATC 619
 QY 580 AGTGACCGAGCTTTCTCGGAGCTGCTGCATCCTTCAACTTCCCTTACCTGCACTGATG 639
 Db 620 AGTGACCGCTGCTTTCAGCTTCTGCGAGAGATTCATTTCTTATCTGCAAGATC 679
 QY 640 TGGCACATCTCTATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699
 Db 680 TGGCATGTGCTCATCAGCATCCTTCCCTTATGCGATGCTCACCATGGCTTGGTGGAT 739
 QY 700 GCTGCCCTCAGAGATTCCTGAGGAGGCTGCTGATCAAGTTCTGCGCCCAATAGAAATGG 759
 Db 740 GCCAACTATGAGATGCCAGGTGAACCCCTCAAAAGTCCGCTACTGGCTCGGGACAGTTGG 799
 QY 760 GCCTTCAATGCTGCTCCCTATGCTGCTCCTCTGTCGCAACCAAGAAATCATCAGTC 816
 Db 800 CC---CGTGGGGCTGCCCTAGTGGAAATCCGGGGTATGACAAAGGACTCTGAGAC 853

RESULT 12

ID AAH48639 standard; DNA; 792 BP.
 XX AAH48639;

XX 21-SEP-2001 (first entry)
 XX Human ceramidase K1 DNA.
 XX Ceramidase; human; K1; antiproliferative; anticancer; anti-eczema;
 KW antipsoriasis; dermatological; ceramide; sphingosine; treatment; ds;
 KW altered cell proliferation; cancer; neurodermatitis; eczema; cosmetic;
 KW permeability barrier; psoriasis; ichthyosis; lamellar ichthyosis ICR2.
 XX Homo sapiens.

XX WO200155408-A1.

XX 02-AUG-2001.

XX 27-JAN-2001; 2001WO-EP000900.

XX 27-JAN-2000; 2000DE-01003293.

PR 09-MAR-2000; 2000DE-01011392.

XX (MEMO-) MEMOREC MEDICAL MOLECULAR RES COLOGNE ST.

XX Hofmann K, Conradt M;

XX WPI; 2001-483256/52.

DR P-PSDB; AAB86363.

XX New ceramidase containing specific structural motif, useful for
 diagnosis, prevention and treatment of ceramidase defects, e.g.
 PT ichthyosis, also in cosmetics.

XX Claim 3; Page 16; 31pp; German.

CC This invention describes novel human ceramidase (I) containing a specific
 CC structural motif which has antiproliferative, anticancer, anti-eczema,
 CC antipsoriasis and dermatological activity. (I) cleave ceramide, resulting
 CC in formation of sphingosine. (II), or the nucleic acid (II) encoding it,
 CC are used for diagnosis, prevention or treatment of diseases associated
 CC with ceramide defects, particularly altered cell proliferation (cancer)
 CC or altered ceramide layers on the skin (neurodermatitis, eczema,
 CC psoriasis), also for targeted modification of the permeability barrier by
 CC ceramidase or its activators, e.g. for transcutaneous delivery of
 CC substances. Specifically they are used for diagnosis of ichthyosis,

CC particularly lamellar ichthyosis ICR2. (I) can also be used in cosmetics.
 CC This sequence encodes human ceramidase K1

XX Sequence 792 BP; 145 A; 241 C; 200 G; 206 T; 0 U; 0 Other;
 SQ

Query Match 22.1%; Score 183.4; DB 4; Length 792;

Best Local Similarity 53.3%; Pred. No. 4.2e-45;

Matches 410; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

QY 40 AGTCCGAGGTGACTGGTGGGAGCACTACACATCGTGTGCTGATTCGCGGAGTTC 99
 Db 25 AGTCCGAGGTGACTGGTGGGAGCACTTCCAGTACTCGGAGCTGGTGGCGGAGTTC 84
 QY 100 TACAACACGATCAGCATGCTTATTTTTCATTTTACCGCCCATCTGCATGTGCTGTTT 159
 Db 85 TACAACACGTTCTCCCAATATCCCTTCTTCACTTCGGGCCACTGATGCTCCTGATG 144
 QY 160 GATGAGTATGCAACATGCTTTGAACAGTGACATCTACTTAACTCTGGAGCTCTTTTGGTTGTA 219
 Db 145 CACCCGATATGCCAGAGCGCTCCGCTACATTTACGTTGCTGGGTCTCTTTCATGATC 204
 QY 220 GTGGGAATTTGGATCCGCTACTTCCATTTTACCTTAGTTTCTTGGGTGAGATGCTTGTAT 279
 Db 205 ATAGGGCTGTTCTTCCATGATTTTCCATGACGCTCAGCTTCTCGGCCAGCTGCTGGAC 264
 QY 280 GAACCTTGCACTGCTTGGGGTCTGATGTGCTTTTGGCCATGTGTTTCCCCAGAGTAT 339
 Db 265 GAGATCGCCATCTCTGGCTCTCTGGGAGTGGCTATAGCATATGGATGCCCCGCTGTAT 324
 QY 340 CTACCAAGATCTTTTCGGAAATGACAGGGGTAGGTTTCAAGGTGGTGGTCACTGCTGCTCT 399
 Db 325 TTCCCTCTCTTCTTGGGGGAAACAGTCCAGTTTATCGGCTGGTCTTTCATCACCACCT 384
 QY 400 GCGGTTACGAGCTGCTTGGCAATTTGCAAGCTTGCATGAAACATCTCTCTGATGACC 459
 Db 385 GTGGTCAGCACCCCTTCTGCTCTTCTCGGGCCCAACGGTCAACGCTACGCCCCCTCAACAGC 444
 QY 460 CTGGGAGTTCCTTGGCACTGCTCATCGCAGAGCTAAAGAGGTGTGACAAACATCGCT 519
 Db 445 ATTGCCCTGCACATCTCTACATCGTGTGCGAGAGTACAGGAGACCACATTAAGAG 504
 QY 520 GTGTTTAAAGCTGGGCTCTTCTCGGGCTCTGGTGACCCCTGGCCCTGTTCTGCTGATC 579
 Db 505 CTTCCGACCTGATTTGAGGTCTCCGTGGTCTTATGGGCTGTGCTCTGACAGCTGATC 564
 QY 580 AGTGACCGAGCTTCTCGGAGCTGCTGTCATCTTCAACTTCCCTTACCTGCACTGCAATG 639
 Db 565 AGTGACCGCTGCTTTCGAGCTTCTGGCAGAGGATTCATTTCTTATCTGCAAGATC 624
 QY 640 TGGCACATCTCTGCTGCTTGTGCTTACCTGGGCTGTGATGCTTGGCTACTTGTAT 699
 Db 625 TGGCATGTGCTCATCAGCATCACCTTCCCTTATGGCATGTGTCACCATGGCTTGGTGGAT 684
 QY 700 GCTGCCCTCAGAGATTCCTGAGCAAGGCCCTTGTGTCATCAAGTTCTTGGCCCAATAGAAATGG 759
 Db 685 GCCAACTATGAGATGCCAGGTGAACCCCTCAAAAGTCCGCTACTGCGCTCGGGACAGTTGG 744
 QY 760 GCCTTCAATGCTGCTCCCTATGCTGCTCCTCTGTCGCAACCAAGAAAT 808
 Db 745 CC---CGTGGGGCTGCCCTACGTTGGAAATCCGGGGTATGACAAAGGACT 790

RESULT 13

ABL24043

ID ABL24043 standard; DNA; 537 BP.

XX ABL24043;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 23602.

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23602.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX Claim 1; SEQ ID NO 23602; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signaling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 537 BP; 102 A; 146 C; 148 G; 141 T; 0 U; 0 Other;
Query Match 17.9%; Score 147.8; DB 4; Length 537;
Best Local Similarity 56.5%; Pred. No. 2.6e-34;
Matches 275; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 19 TGGGACAGCTGCAGGCTGGTAGTCCGAGGTGACCTGGTGGAGGCAACTACCATC 78
Db 49 TGGGAGACCTAGACACCGGAGCTCGCCCTGACTGGTGGAGGCAACTACTTGAT 108
QY 79 GTGCTGCTATGCGCGAGTCTTACACACGATCAGCAATGCTTATTTTCAATTTACCG 138
Db 109 TCGTCCAAATCGCGAGTTCGTGAACACGTTTAGCAACTTCTGTTCATCTACTGCG 168
QY 139 CCCATCTGCATGCTGTTGTTGATGAGTATGCAACATGCTTGAAACAGTGACATCTTA 198
Db 169 CCGCTCTCTTAAATGCTCTTCAAGGAGTACGAGCGCTTGTGACGCCCGGAATCCAGTC 228
QY 199 ATCTGGACTCTTTGGTGTAGTGGGAATGGATCCGCTACTTCCATTTACCTTAGT 258
Db 229 ATCTGGGTGCTGCTCATGCTGGTGGCTGAGTGTGATGATCTCCATGCCACTTGA 288
QY 259 TTCTTGGGTGAGATGCTGTGATGAACCTTGCAGTCCCTTTGGGTCTGTGATGTGCTTGGCC 318
Db 289 CTGATTGGCAGCTGCTGGAGCAACTGGCCATCTCTGGGTCTTCAATGGCGGCTTTTGG 348
QY 319 ATGTGGTTCGCCAAGGATCTACCAAGATCTTTCCGAATGACAGGGGTAGGTTCAAG 378
Db 349 CTCCTTCTATCCGAAGCATATATATCCCAAGTTCGTGAAAAACGATCGMAAACCTTCAGT 408
QY 379 GTGCTGTGCTGAGTGCCTGCTGCGGTACGACGCTGCTGGATTTGTCAAGCTTGGCATC 438
Db 409 TGGCTCATGCTCTTGTGGCGAATTCCTGCGAGGCTTGTGCTGGTGGAGCCCATTTGTT 468
QY 439 AACAAACATCTCTGATGACCCCTGGGAGTTCCTTGGCACTGCACTGCTCATCGCAGAGTA 498
Db 469 AAGCCCTTTGTTCTCATGTTTCAATGATGTCGCGACCATGTTAATGCTCTACAGAGCTG 528

QY 499 AAGAGGT 505
Db 529 CAGAGGT 535
RESULT 14
ABL30432
ID ABL30432 standard; DNA; 4071 BP.
XX ABL30432;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 42769.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX Claim 1; SEQ ID NO 42769; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4071 BP; 1027 A; 938 C; 941 G; 1165 T; 0 U; 0 Other;
Query Match 11.5%; Score 95.6; DB 4; Length 4071;
Best Local Similarity 61.0%; Pred. No. 6.6e-18;
Matches 155; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 512 ACATGCTGTGTTAAGCTGGGCTCTTCTCGGGCTCTGGTGACCTGGCCCTGTCTTCT 571
Db 3224 ACCAGAGGTCTACCGCTGGGATCCGATCGACGACCGTCTGGGCTGTGCGGTGTCT 3283
QY 572 GCTGGATCAGTGACCGAGCTTCTGCGAGCTGCTCATCTTCACTTCCCTTACCTGC 631
Db 3284 GCTGGATCAGTGACCGAGCTTCTGCGAGGCTGCTCTCGATCAATTTCCATACCTGC 3343
QY 632 ACTGCAATGTGGCAATCCTCATCTGCTTGTGCTTACCTGGGCTGTGTATGCTTTGCT 691
Db 3344 ACGGCTTCTGGCACATATTCAATTTTATAGCCGCTACACGGTCTGCTGCTTTTCCCT 3403
QY 692 ACTTTGATGCTGCTCAGATTCCTGAGCAAGGCTGTGCATCAAGTCTGGCCCAATG 751

Db 3404 ACTTCTAGTGGTAATCGGAGCTGCCCCGACGACGCGGCTGTGAAGTACTGGCCAAAGA 3463
 QY 752 AGAAATGGGCTTC 765
 Db 3464 ACGAGTTCGAGTTC 3477

RESULT 15

ABL24042
 ID ABL24042 standard; DNA; 4758 BP.

XX AC ABL24042;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 23599.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical; gene; ds.
 XX OS Drosophila melanogaster.

XX WO200171042-A2.
 XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.
 XX PR 23-MAR-2000; 2000US-0191637P.
 XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WIPI; 2001-656860/75.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

XX PS Claim 1; SEQ ID NO 23599; 21pp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 CC ABBS72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 4758 BP; 1304 A; 981 C; 980 G; 1493 T; 0 U; 0 Other;

Query Match 9.1%; Score 75; DB 4; Length 4758;
 Best Local Similarity 55.6%; Pred. No. 1.4e-11;
 Matches 144; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 108 GATCAGCAATGCTTATTTTTCATTTTACCGCCCATCTGCATGCTGTTGTTGATGAGTA 167
 Db 3287 GTTTAGCAACTTCTGTTTCTTCTACTTGGCGCGCTCTTATATGCTCTTCAAGGAGTA 3346

QY 168 TGCACATGCTTGAACAGTGCATCTACTTAATCTGGACTCTTTTGGTTGTAGTGGAAAT 227
 Db 3347 CGGACGGCTTTGTGACGCGCGGAATCCACGTCATCTGGGTGCTGCTCATCGTGTGGCCT 3406

QY 228 TGGATCGCTTACTTCCATTTTACCTTACTTCTTGGGTGAGTCTTGTATGAACTGC 287
 Db 3407 GAGTTCGATGTACTTCCATGCACTTTGAGTCTGATTTGGCCGCGCTGCTGGACGAACGGC 3466

QY 288 AGTCCTTTGGGTTCTGTGATGTGTGCTTTGGCCATGTGGTTCCCAAGAGGTATCTACCAA 347

Db 3467 CATACTCTGGTCTTTCATGGCGGCTTTTGGCTCTTCTATCCGAGGATACTATCCCAA 3526
 QY 348 GATCTTTCCGAATGACAGG 366
 Db 3527 GTTCGTGAAAAACGATCGG 3545

Search completed: November 20, 2004, 21:28:48
 Job time : 355.193 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:51:27 ; Search time 70.8343 Seconds
(without alignments)
8308.591 Million cell updates/sec

Title: US-10-017-410-3

Perfect score: 828

Sequence: 1 atgggcgcgcgcactggtg.....catcagtcagatcaagtga 828

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A-COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B-COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A-COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B-COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	108.4	13.1	358	4	US-09-621-976-11241
C 2	49.4	6.0	7218	1	Sequence 11241, A
C 3	40.8	4.9	289	3	Sequence 14, Appl
C 4	40.8	4.9	289	3	Sequence 17, Appl
5	40.2	4.9	1063	4	Sequence 875, Appl
6	37.4	4.5	2514	3	Sequence 3, Appl
7	35.8	4.3	1020	4	Sequence 106, Appl
8	35.4	4.3	2340	4	Sequence 67, Appl
9	35.4	4.3	2385	4	Sequence 68, Appl
10	35.4	4.3	2538	4	Sequence 65, Appl
11	35.4	4.3	2553	4	Sequence 66, Appl
12	35.4	4.3	2634	4	Sequence 64, Appl
13	35.4	4.3	4472	4	Sequence 75, Appl
14	35.4	4.3	4766	4	Sequence 73, Appl
15	35.2	4.3	5275	1	Sequence 1, Appl
16	35.2	4.3	5275	1	Sequence 1, Appl
17	35.2	4.3	5275	2	Sequence 1, Appl
18	35.2	4.3	5275	2	Sequence 1, Appl
19	35.2	4.3	5275	3	Sequence 1, Appl
20	35.2	4.3	5275	3	Sequence 1, Appl
21	35.2	4.3	5275	3	Sequence 1, Appl
22	34.8	4.2	505	4	Sequence 15639, A
23	34.8	4.2	1359	4	Sequence 1224, Ap
24	34.6	4.2	505	4	Sequence 15639, A
25	34	4.1	1513	1	Sequence 2, Appl
26	34	4.1	1546	1	Sequence 3, Appl
C 27	34	4.1	2064	1	Sequence 30, Appl

C 28 34 4.1 2064 1 US-08-164-839-32 Sequence 32, Appl

C 29 34 4.1 2064 1 US-08-583-799-30 Sequence 30, Appl

C 30 34 4.1 2064 1 US-08-583-799-32 Sequence 32, Appl

C 31 34 4.1 2148 1 US-08-164-839-69 Sequence 69, Appl

C 32 34 4.1 2148 1 US-08-164-839-71 Sequence 71, Appl

C 33 34 4.1 2148 1 US-08-583-799-69 Sequence 69, Appl

C 34 34 4.1 2148 1 US-08-583-799-71 Sequence 71, Appl

C 35 34 4.1 3172 1 US-08-314-309A-1 Sequence 1, Appl

C 36 33.6 4.1 2016 3 US-09-132-118-1 Sequence 1, Appl

C 37 33.6 4.1 2137 1 US-08-444-005-16 Sequence 16, Appl

C 38 33.6 4.1 2617 3 US-09-161-443-1 Sequence 1, Appl

C 39 33.6 4.1 2617 4 US-09-023-655-914 Sequence 914, App

C 40 33.6 4.1 9687 3 US-09-133-944-2 Sequence 2, Appl

C 41 33.6 4.1 9687 3 US-09-208-827-2 Sequence 2, Appl

C 42 33.6 4.1 9687 4 US-10-043-074-2 Sequence 2, Appl

C 43 33.6 4.1 4403765 3 US-09-103-840A-2 Sequence 1, Appl

C 44 33.6 4.1 4411529 3 US-09-103-840A-1 Sequence 1, Appl

C 45 33.2 4.0 474 4 US-09-621-976-18033 Sequence 18033, A

ALIGNMENTS

RESULT 1

US-09-621-976-11241/c

; Sequence 11241, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 11241

; LENGTH: 358

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-621-976-11241

Query Match 13.1%; Score 108.4; DB 4; Length 358;

Best Local Similarity 99.1%; Pred. No. 2.4e-23;

Matches 109; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGCCCGCACTGGTGGGACCACTGCAGCTGCTAGCTCGGAGTGGACTGGTC 60

Db 111 ATGGGCGCCCGCACTGGTGGGACCACTGCAGCTGCTAGCTCGGAGTGGACTGGTC 52

QY 61 GAGGACAACACTACACCATCGTGCCTGCTATCGCCGAGTTCTACACACGAT 110

Db 51 GAGGACAACACTACACCATCGTGCCTGCTATCGCCGAGTTCTACACACGAT 2

RESULT 2

US-08-232-463-14

; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

```
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/232,463
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/935,313
;; FILING DATE:
;; APPLICATION NUMBER: EP 91 114 300.6
;; FILING DATE: 26-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)836-9300
;; TELEFAX: (703)683-4109
;; TELEX: 899149
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7218 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; CLONE: PTZgpt-Fls
US-08-232-463-14

Query Match 6.0%; Score 49.4; DB 1; Length 7218;
Best Local Similarity 6.1%; Pred. No. 0.00016;
Matches 26; Conservative 218; Mismatches 179; Indels 0; Gaps 0;

QY 369 TAGGTTCAAGTGGTGGTCAAGTCTCTGCTGGGTTACGACGTGCTGSCATTGTGCAA 428
Db 1014 TAGGCTACAGAATTAAATCCGAGCTTGGCTGAGTCCGAGGAGCTTGGGATYYYYYY 1073
QY 429 GCTGCGCATCAACAACATCTCTGTGATGACCCCTGGGAGTCTTGCACATGCTCAT 488
Db 1074 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1133
QY 489 CGCAGAGCTAAGAGGTGTGACACATGCTGTGTTAAGCTGGGCTCTTCTCGGCCT 548
Db 1134 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1193
QY 549 CTGGTGGACCTGGCCCTGTTCTGCTGGATCAGTACCGAGCTTCTCGAGCTGCTGC 608
Db 1194 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1253
QY 609 ATCCTTCAACTCCCTACCTGACATGCTGATGGCAATCCTCATCTGCTTCCCTCA 668
Db 1254 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1313
QY 669 CTGGGCTGTGATGCTTGCCTACTTGTATGCTGCTCAGATTCCTGAGCAAGCCC 728
Db 1314 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1373
QY 729 TGTCAATCAAGTTCTGGCCCAATGAGAAATGGGCTTCAATGTGTCCTATGTGCTC 788
Db 1374 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1433
QY 789 CCT 791
Db 1434 YGT 1436
```

```
RESULT 3
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
```

```
;;
;; GENERAL INFORMATION:
;; APPLICANT: Szostak, Jack W.
;; APPLICANT: Roberts, Richard W.
;; APPLICANT: Liu, Rihe
;; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;; TITLE OF INVENTION: FUSIONS
;; FILE REFERENCE: 00786/350003
;; CURRENT APPLICATION NUMBER: US/09/007,005B
;; CURRENT FILING DATE: 1998-01-14
;; EARLIER APPLICATION NUMBER: 60/035,963
;; EARLIER FILING DATE: 1997-01-27
;; EARLIER APPLICATION NUMBER: 60/064,491
;; EARLIER FILING DATE: 1997-11-06
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 17
;; LENGTH: 289
;; TYPE: RNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Translation template
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)..(289)
;; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match 4.9%; Score 40.8; DB 3; Length 289;
Best Local Similarity 6.9%; Pred. No. 0.011;
Matches 15; Conservative 95; Mismatches 106; Indels 0; Gaps 0;

QY 507 TGACACATGCGTGTGTTAAGCTGGGCTCTTCTCGGCCTCTGCTGGACCTGGCCCT 566
Db 233 YGYCAAGCYTYGYSYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 174
QY 567 GTTCTCTCGATCAGTGACGAGCTTCTCGAGCTGCTGTCATCTTCAACTTCCCTA 626
Db 173 YNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 114
QY 627 CTTGACTGATGCTGACATCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
Db 113 YNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNNYNN 54
QY 687 TGCTACTTGTGCTGCTGCTCAGAGATTCTGAGCA 722
Db 53 YTTGTGTATYATYTYGYAYAYATTAIGTYA 18

RESULT 4
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
;; GENERAL INFORMATION:
;; APPLICANT: Szostak, Jack W.
;; APPLICANT: Roberts, Richard W.
;; APPLICANT: Liu, Rihe
;; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;; TITLE OF INVENTION: FUSIONS
;; FILE REFERENCE: 00786/350007
;; CURRENT APPLICATION NUMBER: US/09/244,796
;; CURRENT FILING DATE: 1999-02-05
;; EARLIER APPLICATION NUMBER: 60/035,963
;; EARLIER FILING DATE: 1997-01-27
;; EARLIER APPLICATION NUMBER: 60/064,491
;; EARLIER FILING DATE: 1997-11-06
;; EARLIER APPLICATION NUMBER: 09/007,005
;; EARLIER FILING DATE: 1998-01-14
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 17
;; LENGTH: 289
;; TYPE: RNA
;; ORGANISM: Artificial Sequence
```

```
;
;
; FEATURE:
; OTHER INFORMATION: Translation template
;
; NAME/KEY: misc.feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match      4.9%; Score 40.8; DB 3; Length 289;
Best Local Similarity 6.9%; Pred. No. 0.011;
Matches 15; Conservative 95; Mismatches 106; Indels 0; Gaps 0;

QY 507 TGACACATCGTGTGTTAAGCTGGGCTCTTCGCGGCTCTGGTGACCCCTGGCCCT 566
Db 233 YGYAYGCTGYGYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 174
QY 567 GTTCTGCTGGATCAGTGACCGAGCTTTCTGCGAGCTGCTGCATCCTTCAACTTCCCTCA 626
Db 173 YNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 114
QY 627 CTGCACTGATGGGACATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
Db 113 YNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYN 54
QY 687 TGCTACTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
Db 53 YTYTYGYAYAYTYTYGYAYAYAYAYAYAYAYAYAYAYAYAYAYAYAYAY 18

RESULT 5
US-09-799-451-875
; Sequence 875, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 875
; LENGTH: 1063
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(802)
US-09-799-451-875

Query Match      4.9%; Score 40.2; DB 4; Length 1063;
Best Local Similarity 57.6%; Pred. No. 0.035;
Matches 72; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 7 GCCCGCACTGGTGGGACCACTGCAGGCTGGTAGCTCGGAGGTGGACTGGTGGAGGAC 66
```

```
Db 11 GCCCGCACTGGGAGGGCTACTGGGGCCCCACGACCTCCACGCTGGACTGGTGGAGGAG 70
QY 67 AACTACACCATCGTGCCTGCTATCGCCGAGTTCTACAAACAGATCAGCAATGCTCTTATT 126
Db 71 AACTACTCCGTGACCTGGTACATCGCGAGTCTTGGAAATACAGTGAAGTAACCTGATCATG 130
QY 127 TTCAT 131
Db 131 ATTAT 135

RESULT 6
US-09-144-914-3
; Sequence 3, Application US/09144914
; Patent No. 6309855
; GENERAL INFORMATION:
; APPLICANT: Duprat, Fabrice
; APPLICANT: Lesage, Florian
; APPLICANT: Fink, Michel
; APPLICANT: Lazdunski, Michel
; TITLE OF INVENTION: FAMILY OF MAMMALIAN POTASSIUM CHANNELS, THEIR CLONING
; FILE REFERENCE: 989.6705CIP
; CURRENT APPLICATION NUMBER: US/09/144,914
; CURRENT FILING DATE: 1998-09-01
; EARLIER APPLICATION NUMBER: 08/749,816
; EARLIER FILING DATE: 1996-11-15
; EARLIER APPLICATION NUMBER: 60/095,234
; EARLIER FILING DATE: 1998-08-04
; EARLIER APPLICATION NUMBER: FR 96/01565
; EARLIER FILING DATE: 1996-02-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (126)...(1307)
; FEATURE:
; OTHER INFORMATION: TASK
US-09-144-914-3

Query Match      4.5%; Score 37.4; DB 3; Length 2514;
Best Local Similarity 50.9%; Pred. No. 0.41;
Matches 89; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 499 AAGAGGTGTGACAAACATGCGTGTGTTTAAAGCTGGGCTCTTCTGGGCTCTGCTGGACC 558
Db 570 ATGCGGCGCGGAGGTGTCATGGCCAAACATGCTGCTCATCGGCTTCTTCTGCTGCATC 629
QY 559 CTGGCCCTGTTCTGCTGGATCAGTACCGAGCTTCTGCGAGCTGCTGCTCATCTTCAAC 618
Db 630 AGCAGCGTGTGCATCGCGCGCGCTTCTTCCCACTACGAGCACTGACCTTCTTCCAG 689
QY 619 TTCCCTTACCTGCACTGCACTGCGCATGCGCACATCTCATCTGCTGCTGCTGCTGCTGCT 673
Db 690 GCCTACTACTGCTTCTCATCCCTTACCACCATCGGCTTCCGCGACTACGTTGG 744

RESULT 7
US-09-489-039A-106
; Sequence 106, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLBBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
```

Query Match 4.3%; Score 35.4; DB 4; Length 2538;
Best Local Similarity 57.8%; Pred. No. 1.7;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 73
LENGTH: 4766
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: gp160.modUSA.gag.modSF2
US-09-475-515-73

Query Match 4.3%; Score 35.4; DB 4; Length 4766;
Best Local Similarity 57.8%; Pred. No. 2.4;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 12 GCACGTGTTGGGACCGAGCTGCGAGGTGCTGCTGAGGTGAGTGGTGGGAGGACACTA 71
Db 2427 GTACTGTGGAACCTGCTGAGTACTGGAGCCGAGGACTGAAGAGAGCCCGGTGACCT 2486

QY 72 CACCATGTCGCTGCTATCGCCGAGTTCTACAAACAGCATCAGCAATGTC 120
Db 2487 GTTCAACGCCACCGCCATCGCGTGGCGGAGGACCGACCGCATCATC 2535

RESULT 15
US-08-485-588-1
Sequence 1, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994

Query Match 4.3%; Score 35.2; DB 1; Length 5275;
Best Local Similarity 50.6%; Pred. No. 3;
Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 483 GCTCATCGCAGAGCTAAAGAGGTGTGACAAACATGCGTGTGTTAAAGCTGGGCGCTTCTC 542
Db 2437 GCCATGCTCAAGGCCACCAACCGGAGCTCTCCTATCTCTTCTCTTCTCTCTCTG 2496

QY 543 GGGCCTTGTGTGACCCCTGGCCCTGTTCTGTCGATCAGTACCGAGCTTTCTCGAGCT 602
Db 2497 CTGCTTCTCCAGCTCCCTGTTCTTCTATCGGGGAGCCAGAGCTGGAGCTGCCGCTCG 2556

QY 603 GCTGTCTATCTTCAACTTCCCTACCTGCACTGCTGTCATGTCATCTCT 650
Db 2557 CCAGCGCGCCTTTGGCATCAGCTTCGTGCTCTGTCATCTCTGTCATCTCT 2604

Search completed: November 21, 2004, 07:04:51
Job time : 75.8343 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:53:12 ; Search time 386.775 Seconds
(without alignments)
11567.962 Million cell updates/sec

Title: US-10-017-410-3

Perfect score: 828

Sequence: 1 atggcgcccgccactggtg.....catcagtcagatcacgtga 828

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	828	100.0	828	13	US-10-017-410-3
2	813.6	98.3	4202	16	US-10-188-832-21
3	813.6	98.3	4212	15	US-10-295-027-151
4	761	91.9	822	15	US-10-182-447-6
5	664.8	80.3	4175	13	US-10-017-410-1
6	350.6	42.3	744	10	US-09-945-527-48
7	183.4	22.1	792	15	US-10-182-447-4
8	52.2	6.3	35425	15	US-10-017-161-2429
9	52.2	6.3	35425	15	US-10-292-798-2069
10	42.2	5.1	3034	18	US-10-425-115-77895
11	41.8	5.0	1604	17	US-10-437-963-77989
12	41	5.0	637	18	US-10-425-115-146588

c	13	40.2	4.9	284	15	US-10-002-631C-17	Sequence 17, Appl
	14	40.2	4.9	636	16	US-10-264-237-615	Sequence 615, App
	15	40.2	4.9	801	15	US-10-182-447-5	Sequence 5, Appli
	16	40.2	4.9	1063	16	US-10-302-172-875	Sequence 875, App
	17	39.2	4.7	598	17	US-10-767-701-20737	Sequence 20737, A
	18	39	4.7	1194	16	US-10-425-114-32295	Sequence 32295, A
	19	39	4.7	1429	18	US-10-425-115-107231	Sequence 107231, A
	20	39	4.7	1471	17	US-10-767-701-12261	Sequence 12261, A
	21	39	4.7	1547	17	US-10-029-386-27071	Sequence 27071, A
	22	38.8	4.7	332	15	US-10-029-386-27071	Sequence 27071, A
	23	38.8	4.7	579	15	US-10-029-386-27071	Sequence 27071, A
	24	38.8	4.7	744802	15	US-10-252-798-13371	Sequence 13371, A
	25	38.6	4.7	357	18	US-10-425-115-24309	Sequence 24309, A
	26	38	4.6	158405	14	US-10-175-523-86	Sequence 86, Appl
	27	37.4	4.5	854	9	US-09-746-491-55	Sequence 55, Appl
	28	37.4	4.5	2514	9	US-09-939-484-3	Sequence 3, Appli
	29	37.4	4.5	2514	9	US-09-939-484-3	Sequence 3, Appli
	30	37.4	4.5	2590	15	US-10-146-733-7	Sequence 7, Appli
	31	37.2	4.5	700	16	US-10-260-238-3722	Sequence 3722, Ap
	32	37.2	4.5	25576	10	US-09-764-891-8306	Sequence 8306, Ap
	33	37	4.5	851	18	US-10-425-115-53388	Sequence 53388, A
	34	36.8	4.4	3657	16	US-10-388-934-165	Sequence 165, App
	35	36.8	4.4	3657	16	US-10-191-803-311	Sequence 311, App
	36	36.6	4.4	607	14	US-10-123-155-344	Sequence 344, App
	37	36.6	4.4	607	15	US-10-146-731-344	Sequence 344, App
	38	36.6	4.4	607	15	US-10-140-472-344	Sequence 344, App
	39	36.6	4.4	607	15	US-10-141-761-344	Sequence 344, App
	40	36.6	4.4	607	15	US-10-142-885-344	Sequence 344, App
	41	36.6	4.4	607	15	US-10-158-790-344	Sequence 344, App
	42	36.6	4.4	607	15	US-10-137-871-344	Sequence 344, App
	43	36.6	4.4	607	15	US-10-140-923-344	Sequence 344, App
	44	36.6	4.4	607	15	US-10-141-756-344	Sequence 344, App
	45	36.6	4.4	607	15	US-10-141-759-344	Sequence 344, App

ALIGNMENTS

RESULT 1
US-10-017-410-3
; Sequence 3, Application US/10017410
; Publication No. US20020115094A1
; GENERAL INFORMATION:
; APPLICANT: Farnham, Peggy J
; APPLICANT: Graveel, Carrie R
; TITLE OF INVENTION: Polynucleotide Differentially Expressed in Liver Cancer
; FILE REFERENCE: 960296.97401
; CURRENT APPLICATION NUMBER: US/10/017,410
; CURRENT FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(825)
US-10-017-410-3

Query Match 100.0%; Score 828; DB 13; Length 828;
Best Local Similarity 100.0%; Pred. No. 2.2e-253;
Matches 828; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGCGCCCGCCACTGTTGGGACCCAGCTCGAGCTGGTAGCTCGGAGTGGAGCTGTGC	60
Db	1	ATGGCGCCCGCCACTGTTGGGACCCAGCTCGAGCTGGTAGCTCGGAGTGGAGCTGTGC	60
Qy	61	GAGGACAACATACACCATCGTCTGCTATCGCGAGTTCTACAAACGATCAGCAATGTC	120
Db	61	GAGGACAACATACACCATCGTCTGCTATCGCGAGTTCTACAAACGATCAGCAATGTC	120
Qy	121	TTATTTTTCATTATACCGCCCATCTGCATGTGCTTTTGTGATGATGATGATGATGATGTC	180

121	TTATTTTTCATTTCACGCCCATCTGCATGTGCTTGTTTGATCAGTATGCAACATCGTTG	180	DB
181	AACAGTGACATCTACTTAACTCTGACATCTTTTGGTGTGTAGTGGAAATGGATCCGTCTAC	240	QY
191	AACAGTGACATCTACTTAACTCTGACATCTTTTGGTGTGTAGTGGAAATGGATCCGTCTAC	240	DB
241	TTCCATTTTACCCCTTAGTTCCTTGGGTGAGATGCTTGATGAACTTCGACGTCCTTTGGGTT	300	QY
241	TTCCATTTTACCCCTTAGTTCCTTGGGTGAGATGCTTGATGAACTTCGACGTCCTTTGGGTT	300	DB
301	CTGATGTGTCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTTCGGAAT	360	QY
301	CTGATGTGTCTTTGGCCATGTGTTCCCGAAGGTATCTACCAAGATCTTTTCGGAAT	360	DB
361	GACAGGGGTAGTGTCAAGTGTGTGTCAGTGTCTGTCTGGTGTACGAGTCGCTTGGA	420	QY
361	GACAGGGGTAGTGTCAAGTGTGTGTCAGTGTCTGTCTGGTGTACGAGTCGCTTGGA	420	DB
421	TTTGTCAAGCCTGCCATCAACAACATCTCTCTGTAAGCCCTGGAGTTCCTTGACATGCA	480	QY
421	TTTGTCAAGCCTGCCATCAACAACATCTCTCTGTAAGCCCTGGAGTTCCTTGACATGCA	480	DB
481	CTGCTCATCGCAGAGCTAAGAGGTGTGACAACATCGTGTGTTTAAGCTGGGCCCTTTC	540	QY
481	CTGCTCATCGCAGAGCTAAGAGGTGTGACAACATCGTGTGTTTAAGCTGGGCCCTTTC	540	DB
541	TCGGGCCCTCTGTGTGACCTCGGCCCTGTTCTGTGGATCAGTGACCCGAGCTTTCTCGAG	600	QY
541	TCGGGCCCTCTGTGTGACCTCGGCCCTGTTCTGTGGATCAGTGACCCGAGCTTTCTCGAG	600	DB
601	CTGCTGTATCCTTCAACTTCCTCCCTACCTGCACTGTGCATGTGGCACATCTCATCTGCTT	660	QY
601	CTGCTGTATCCTTCAACTTCCTCCCTACCTGCACTGTGCATGTGGCACATCTCATCTGCTT	660	DB
661	GCTGCTACCTGGCTGTGTATGCTTTTGGCTACTTTTGATGCTGCTCAGAGATTCCTGAG	720	QY
661	GCTGCTACCTGGCTGTGTATGCTTTTGGCTACTTTTGATGCTGCTCAGAGATTCCTGAG	720	DB
721	CAAGGCCCTGTGCATCAAGTTCTGGGCCCAATGAGAAATGGGCTTTCATTGTGTCCCCCTAT	780	QY
721	CAAGGCCCTGTGCATCAAGTTCTGGGCCCAATGAGAAATGGGCTTTCATTGTGTCCCCCTAT	780	DB
781	GTGTGCTCCTCTGTGTGCCAAACAAGAAATCATCAGTCAAGATCACTGA	828	QY
781	GTGTGCTCCTCTGTGTGCCAAACAAGAAATCATCAGTCAAGATCACTGA	828	DB

RESULT 2

```

US-10-188-832-21
; Sequence 21, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 4202
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-21

Query Match      98.3%; Score 813.6; DB 16; Length 4202;
Best Local Similarity 98.9%; Pred. No. 2e-248; 9; Indels 0; Gaps 0;
Matches 819; Conservative 0; Mismatches 9;

Qy 1  ATGGGGCCCCGACACTGGTGGGACCAAGCTGACGGCTGGTAGCTCGGAGGTGGACTGGTGC 60
Db 77  ATGGGGCCCCGACACTGGTGGGACCAAGCTGACGGCTGGTAGCTCGGAGGTGGACTGGTGC 136
Qy 61  GAGGACAACTACACCATCGTGCCTGCTATCGCCGAGTTCTAACACAGCATCAGCAATGTC 120
Db 137  GAGGACAACTACACCATCGTGCCTGCTATCGCCGAGTTCTAACACAGCATCAGCAATGTC 196
Qy 121  TTAATTTTCAATTTTACGCCCCATCTGCATGTGCTGTTTGATCAGATATGCAACATGCTTG 180
Db 197  TTAATTTTCAATTTACGCCCCATCTGCATGTGCTGTTTGATCAGATATGCAACATGCTTG 256
Qy 181  AACAGTGCATCTACTTAATCTGGACTCTTTTGGTTGTTAGTGGGAATTGATCGCTCTAC 240
Db 257  AACAGTGGCATCTACTTAATCTGGACTCTTTTGGTTGTTAGTGGGAATTGATCGCTCTAC 316
Qy 241  TTCCATTTTACCCTTAGTTTCTTGGGTTCAGATGCTTGATGAACCTGCAGTCCCTTTGGGTT 300
Db 317  TTCATGTCAAACCCCTTAGTTTCTTGGGTTCAGATGCTTGATGAACCTGCAGTCCCTTTGGGTT 376
Qy 301  CTGATGTGTCCTTTGGCCATGTGGTCTCCCCAGAAAGGTATCTACCAAAGATCTTTCGGAAT 360
Db 377  CTGATGTGTCCTTTGGCCATGTGGTCTCCCCAGAAAGGTATCTACCAAAGATCTTTCGGAAT 436
Qy 361  GACAGGGGTAGGTTCAAGGTGGTGGTCAGTGTCTGTCTGGGTACGAGTGCCTGGCA 420
Db 437  GACCGGGTAGGTTCAAGGTGGTGGTCAGTGTCTGTCTGGGTACGAGTGCCTGGCA 496
Qy 421  TTTGTCAAGCTGTCATCAACAAATCTCTGTATGACCTTGGGAGTTCCTTGCACGTCA 480
Db 497  TTTGTCAAGCTGTCATCAACAAATCTCTGTATGACCTTGGGAGTTCCTTGCACGTCA 556
Qy 481  CTGCTCATCGCAGAGCTAAAGAGGTGTGCAACATGGGTGTGTTTAAAGTGGGCTCTTTC 540
Db 557  CTGCTCATCGCAGAGCTAAAGAGGTGTGCAACATGGGTGTGTTTAAAGTGGGCTCTTTC 616
Qy 541  TCGGGCCCTGTGGTGGACCTGGCCCTGTTCTGCTGGATCAGTGACCCAGCTTCTTCGGAG 600
Db 617  TCGGGCCCTGTGGTGGACCTGGCCCTGTTCTGCTGGATCAGTGACCCAGCTTCTTCGGAG 676
Qy 601  CTGCTGTGATCTCTTCAACTTCCCCCTACTGCACTGTCATGTGGCACATCTTCATCTGCCTT 660
Db 677  CTGCTGTGATCTCTTCAACTTCCCCCTACTGCACTGTCATGTGGCACATCTTCATCTGCCTT 736
Qy 661  GCTGCCCTACCTGGGCTGTGTATGCTTTGGCTTACTTTGATGTGCTCTCAGAGATTCCTGAG 720
Db 737  GCTGCCCTACCTGGGCTGTGTATGCTTTGGCTTACTTTGATGTGCTCTCAGAGATTCCTGAG 796
Qy 721  CAAAGGCCCTGTGCATCAAGTTCTGGCCCAATGAGAAATGGGGCTTTCATTGGTGTCCCTTAT 780
Db 797  CAAAGGCCCTGTGCATCAAGTTCTGGCCCAATGAGAAATGGGGCTTTCATTGGTGTCCCTTAT 856
Qy 781  GTGTCCCTCTCTGTGTGGCCAAACAGAAATCATCAGTCAAGATCACGTGA 828
Db 857  GTGTCCCTCTCTGTGTGGCCAAACAGAAATCATCAGTCAAGATCACGTGA 904

```

RESULT 3
US-10-295-027-151
; Sequence 151, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:

APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 151
LENGTH: 4212
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-151

Query Match 98.3%; Score 813.6; DB 15; Length 4212;
Best Local Similarity 98.9%; Pred. No. 2e-248;
Matches 819; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ATGGGCGCCCGCACTGGTGGGACCACTGCAGGCTGTAGCTCGGAGGTGGACTGGTC 60
Db 87 ATGGGCGCCCGCACTGGTGGGACCACTGCAGGCTGTAGCTCGGAGGTGGACTGGTC 146

Qy 61 GAGGACAACTACACCACTCGTGCCTGTCTACCGAGTCTTACAAACGATCAGCAATGTC 120
Db 147 GAGGACAACTACACCACTCGTGCCTGTCTACCGAGTCTTACAAACGATCAGCAATGTC 206

Qy 121 TTATTTTTCATTTACCGCCCACTGCATGTGCTTTGATGAGTATGCAACATGCTT 180
Db 207 TTATTTTTCATTTACCGCCCACTGCATGTGCTTTGATGAGTATGCAACATGCTT 266

Qy 181 AACAGTCACATCTACTTAATCTGACCTTTTGGTGTAGTGGAAATGGATCGCTTAC 240
Db 267 AACAGTCACATCTACTTAATCTGACCTTTTGGTGTAGTGGAAATGGATCGCTTAC 326

Qy 241 TTCATTTTACCTTTAGTTTCTTGGGTTCAGATGCTTTCAGTTCCTTTTGGGT 300
Db 327 TTCATTCACACCTTAGTTTCTTGGGTTCAGATGCTTTCAGTTCCTTTTGGGT 386

Qy 301 CTGATGTGCTTTGGCCATGTGTTCCCGAAGAGTATCTACCAAGATCTTTTCGGAAT 360
Db 387 CTGATGTGCTTTGGCCATGTGTTCCCGAAGAGTATCTACCAAGATCTTTTCGGAAT 446

Qy 361 GACAGGGGTAGTTCAAGTGGTGGTTCAGTGTCTCTGTCTCGGTTACGAGTGCCTGGCA 420

Db 447 GACCGGGGTAGTTCAAGGTGGTGGTCTCAGTGTCTCTGCTGCGGTACGACGTGCTGGCA 506

Qy 421 TTTGTCAAGCCTGCATCAACAAATCTCTCTGATGACCCCTGGGAGTTCTTGGCACTGCA 480

Db 507 TTTGTCAAGCCTGCATCAACAAATCTCTCTGATGACCCCTGGGAGTTCTTGGCACTGCA 566

Qy 481 CTGCTCATTCGAGAGCTAAAGAGGTGTGACAAATGCGTGTGTTTAAAGCTGGGCTCTTC 540

Db 567 CTGCTCATTCGAGAGCTAAAGAGGTGTGACAAATGCGTGTGTTTAAAGCTGGGCTCTTC 626

Qy 541 TCGGGCCTCTGGTGGACCCCTGGCCCTGTCTGCTGGATCACTGACCGAGCTTTCGCGAG 600

Db 627 TCGGGCCTCTGGTGGACCCCTGGCCCTGTCTGCTGGATCACTGACCGAGCTTTCGCGAG 686

Qy 601 CTGCTGTCTCTCTCAACTTCCCTCTACCTGACATGATGTCGACATCTCTCATCTGCTT 660

Db 687 CTGCTGTCTCTCTCAACTTCCCTCTACCTGACATGATGTCGACATCTCTCATCTGCTT 746

Qy 661 GCTCCCTACCTGGCTGTGATGCTTTGCTCTTCTTGTGCTGCTCAGAGATCTCTGAG 720

Db 747 GCTCCCTACCTGGCTGTGATGCTTTGCTCTTCTTGTGCTGCTCAGAGATCTCTGAG 806

Qy 721 CAAGGCGCTGTCTCAAGTTCTGCGCCCAATGAGAAATGGGCTTTCATTTGGTCCCTAT 780

Db 807 CAAGGCGCTGTCTCAAGTTCTGCGCCCAATGAGAAATGGGCTTTCATTTGGTCCCTAT 866

Qy 781 GTGTCCCTCTCTGTGTCGCAACAAAGAAATCATCAGTCAAGATCAGTGA 828

Db 867 GTGTCCCTCTCTGTGTCGCAACAAAGAAATCATCAGTCAAGATCAGTGA 914

RESULT 4
US-10-182-447-6
; Sequence 6, Application US/10182447
; Publication No. US20030185814A1
; GENERAL INFORMATION:
; APPLICANT: HOFMANN, Kay
; APPLICANT: RADT, Marcus
; TITLE OF INVENTION: CERAMIDASE
; FILE REFERENCE: P68055USO
; CURRENT APPLICATION NUMBER: US/10/182,447
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: PCT/EP01/00900
; PRIOR FILING DATE: 2001-01-27
; PRIOR APPLICATION NUMBER: DE 10003293.1
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: DE 10011392.3
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-447-6

Query Match 91.9%; Score 761; DB 15; Length 822;
Best Local Similarity 96.0%; Pred. No. 5.4e-232;
Matches 792; Conservative 0; Mismatches 30; Indels 3; Gaps 1;

Qy 1 ATGGGCGCCCGCACTGGTGGGACCACTGCAGGCTGTAGCTCGGAGGTGGACTGGTC 60
Db 1 ATGGGCGCCCGCACTGGTGGGACCACTGCAGGCTGTAGCTCGGAGGTGGACTGGTC 60

Qy 61 GAGGACAACTACACCACTGTGCTCTATCGCGAGTCTTACAAACGATCAGCAATGTC 120
Db 61 GAGGACAACTACACCACTGTGCTCTATCGCGAGTCTTATTAACATGATCAGCAATGTC 120

Qy 121 TTATTTTTCATTTTACCGCCCACTGCATGTGCTTTGATGAGTATGCAACATGCTT 180
Db 121 TTATTTTTCATTTTACCGCCCACTGCATGTGCTTTGATGAGTATGCAACATGCTT 180


```
US-10-437-963-77989/c
; Sequence 77989, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77989
; LENGTH: 1604
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1604)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77836C.1
US-10-437-963-77989

Query Match          5.0%; Score 41.8; DB 17; Length 1604;
Best Local Similarity 52.6%; Pred. No. 0.035;
Matches 91; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 513 CATCGGTGTGTTAAGCTGGCCCTCTTCGGGCTCTGTGGGACCTGGCCCTGTTCTG 572
Db 1375 CATACGGTGTGAGGTCATCTCTCTTCAGATCGGTCGAGCTCCGGCGGATCAG 1316
Qy 573 CTGATCAGTGACCGAGCTTFTCTCGAGCTCTGTGCATCTTCAACTTCCCTACCTGCA 632
Db 1315 CCGCATCTTCGGCGCGCTCTCGCGAGCTCTGACGACGTCAACGTCAGCCGACCG 1256
Qy 633 CTGCATGTGGCACATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 685
Db 1255 CTCCTTCATCCCTCCGACGTAAGCTGCCTCCCAACCTTGGATCCGTCACCT 1203

RESULT 12
US-10-425-115-146588
; Sequence 146588, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 146588
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_65204C.1
US-10-425-115-146588

Query Match          5.0%; Score 41; DB 18; Length 637;
Best Local Similarity 52.7%; Pred. No. 0.039;
Matches 89; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

US-10-437-963-77989/c
; Sequence 77989, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77989
; LENGTH: 1604
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1604)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77836C.1
US-10-437-963-77989

Query Match          5.0%; Score 41.8; DB 17; Length 1604;
Best Local Similarity 52.6%; Pred. No. 0.035;
Matches 91; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 513 CATCGGTGTGTTAAGCTGGCCCTCTTCGGGCTCTGTGGGACCTGGCCCTGTTCTG 572
Db 1375 CATACGGTGTGAGGTCATCTCTCTTCAGATCGGTCGAGCTCCGGCGGATCAG 1316
Qy 573 CTGATCAGTGACCGAGCTTFTCTCGAGCTCTGTGCATCTTCAACTTCCCTACCTGCA 632
Db 1315 CCGCATCTTCGGCGCGCTCTCGCGAGCTCTGACGACGTCAACGTCAGCCGACCG 1256
Qy 633 CTGCATGTGGCACATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 685
Db 1255 CTCCTTCATCCCTCCGACGTAAGCTGCCTCCCAACCTTGGATCCGTCACCT 1203

RESULT 12
US-10-425-115-146588
; Sequence 146588, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 146588
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_65204C.1
US-10-425-115-146588

Query Match          5.0%; Score 41; DB 18; Length 637;
Best Local Similarity 52.7%; Pred. No. 0.039;
Matches 89; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

US-10-002-631C-17/c
; Sequence 17, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-002-631C-17

Query Match          4.9%; Score 40.2; DB 15; Length 284;
Best Local Similarity 57.6%; Pred. No. 0.047;
Matches 72; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 7 GCCCGCGACCTGTGTGGGACCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 66
Db 222 GCCCGCGACCGAGAGGGCTACTGGGGCCCGACGACCTCCACGCTGGACTGGTGGAGG 163
Qy 67 AACTACACCATCGTGCCTGCTATCGCGAGTCTTACACAGCATCAGCAATGCTTATTT 126
Db 162 AACTACTCCGTGACCTGTGTACATCCCGAGTCTTGAATACAGTGAGTAACCTGATCATG 103
Qy 127 TTCTAT 131
Db 102 ATTAT 98

RESULT 14
US-10-264-237-615
; Sequence 615, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 615
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
Qy 527 AGCTGGGCCCTCTTCTCGGGCTCTGTGGGACCTGGCCCTGTTCTGCTGGATCATGACC 586
Db 88 AGCGCGCGCGAGGCGGCTCGCGGGGACCAACGCCACGTCCTCAAGCTCACACACC 147
Qy 587 GAGCTTTCTGCGAGCTGTGTCATCTTCAACTTCCCTTACCTGCACTGCAATGTGGACA 646
Db 148 GCGCTCTCTCGAGCTGTGCTGCTGAGCAGCTTCAAGAACCTGGAGCACTCGACCTCG 207
Qy 647 TCCTCATCTGCTTGTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTTTCCTACTT 695
Db 208 GCTTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
```

```
RESULT 13
US-10-002-631C-17/c
; Sequence 17, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-002-631C-17
```

```
Query Match          4.9%; Score 40.2; DB 15; Length 284;
Best Local Similarity 57.6%; Pred. No. 0.047;
Matches 72; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 7 GCCCGCGACCTGTGTGGGACCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 66
Db 222 GCCCGCGACCGAGAGGGCTACTGGGGCCCGACGACCTCCACGCTGGACTGGTGGAGG 163
Qy 67 AACTACACCATCGTGCCTGCTATCGCGAGTCTTACACAGCATCAGCAATGCTTATTT 126
Db 162 AACTACTCCGTGACCTGTGTACATCCCGAGTCTTGAATACAGTGAGTAACCTGATCATG 103
Qy 127 TTCTAT 131
Db 102 ATTAT 98
```

```
RESULT 14
US-10-264-237-615
; Sequence 615, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 615
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

Job time : 389.775 secs

```
; NAME/KEY: misc feature
; LOCATION: (560)..(560)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (631)..(631)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-615
```

```
Query Match 4.9%; Score 40.2; DB 16; Length 636;
Best Local Similarity 57.6%; Pred. No. 0.07;
Matches 72; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 7 GCCCGGCACTGGTGGGACCAAGCTGAGCTGGAGGTGGAGTGGTGGAGGAC 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 GCCGGGACCGAGAGGGCTACTGGGGCCCAAGACCTCCACGCTGGTGGAGGAG 119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 67 AACTACACCATCGTCTGCTATCGCGAGTTCTTACAACAGATCAGCAATGCTTATTT 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 AACTACTCCGTGACCTGGTACATCGCGAGTTCTTGAATACAGTGAGTAACCTGATCATG 179
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 127 TTCAT 131
   |||
Db 180 ATTAT 184
   |||
```

```
RESULT 15
US-10-182-447-5
; Sequence 5, Application US/10182447
; Publication No. US20030185814A1
; GENERAL INFORMATION:
; APPLICANT: HOFMANN, Kay
; TITLE OF INVENTION: CERAMIDASE
; FILE REFERENCE: P68055US0
; CURRENT APPLICATION NUMBER: US/10/182,447
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: PCT/EP01/00900
; PRIOR FILING DATE: 2001-01-27
; PRIOR APPLICATION NUMBER: DE 10003293.1
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: DE 10011392.3
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-447-5
```

```
Query Match 4.9%; Score 40.2; DB 15; Length 801;
Best Local Similarity 57.6%; Pred. No. 0.079;
Matches 72; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 7 GCCCGGCACTGGTGGGACCAAGCTGAGCTGGAGGTGGAGTGGTGGAGGAC 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 GCCGGGACCGAGAGGGCTACTGGGGCCCAAGACCTCCACGCTGGTGGAGGAG 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 67 AACTACACCATCGTCTGCTATCGCGAGTTCTTACAACAGATCAGCAATGCTTATTT 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 AACTACTCCGTGACCTGGTACATCGCGAGTTCTTGAATACAGTGAGTAACCTGATCATG 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 127 TTCAT 131
   |||
Db 130 ATTAT 134
   |||
```

Search completed: November 21, 2004, 07:44:02

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2004, 20:50:47 ; Search time 2284.57 Seconds
(without alignments)
13206.895 Million cell updates/sec

Title: US-10-017-410-3

Perfect score: 828
Sequence: 1 atggggccccgcactgggtg.....catcagtcacagatcacgtga 828

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hlc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	671.4	81.1	681	9 AY401889	AY401889 Homo sapi
2	650.2	79.7	681	9 AY401890	AY401890 Pan trogl
3	593.6	71.7	1022	5 BU511164	BU511164 AGENCOURT
4	535.4	64.7	681	9 AY401891	AY401891 Mus muscu
5	512.4	61.9	868	6 CA976684	CA976684 AGENCOURT
6	512.4	61.9	1173	3 AK085306	AK085306 Mus muscu
7	480	58.0	797	6 CA463294	CA463294 AGENCOURT
8	454.4	54.9	565	6 CF115220	CF115220 Shultzomi
9	447.4	54.0	449	5 BX646596	BX646596 DKFZp781B
10	427.2	51.6	633	2 BB660847	BB660847
11	416.2	50.3	498	4 BT848265	BT848265 470659 MA
12	414.8	50.1	1527	3 AF370405	AF370405 Homo sapi
13	403	48.7	698	5 BU234223	BU234223 603792086
14	381	46.0	781	5 BU232554	BU232554 603409105
15	372.6	45.0	548	6 CF169369	CF169369 B0812G07-
16	361.4	43.6	847	5 BU220481	BU220481 603107516
17	348.8	42.1	678	5 BU203269	BU203269 604153603
18	343.6	41.5	937	5 EX842808	EX842808 BX842808
19	341.4	41.2	432	6 CF169808	CF169808 B0818D08-
20	340.6	41.1	648	4 BJ060795	BJ060795
21	329.4	39.8	454	6 CB784796	CB784796 AMGNNUC:T
22	325.2	39.3	427	1 AA900336	AA900336 UI-R-EQ-C
23	324.2	39.2	401	2 BF554219	BF554219 UI-R-CO-h
24	306.4	37.0	434	6 CB758986	CB758986 AMGNNUC:M

25	300.2	36.3	622	6	CB723138	CB723138 UI-M-GH0-
26	299.8	36.2	480	6	CB728612	CB728612 AMGNNUC:M
27	299.4	36.2	399	6	CB706492	CB706492 AMGNNUC:M
28	285.2	34.4	803	7	CF593785	CF593785 AGENCOURT
29	256.6	31.0	295	2	BE668106	BE668106 156540 MA
30	242.4	29.3	365	5	BY117228	BY117228 BY117228
31	240.2	29.0	357	5	BY168309	BY168309 BY168309
c	239	28.9	485	2	BF549345	BF549345 UI-R-A0-a
33	227	27.4	423	6	CA967276	CA967276 CCLX06a22
34	225.4	27.1	646	6	CA969385	CA969385 CCLX06a22
35	224.2	27.1	815	5	BX623644	BX623644 BX623644
36	212.6	25.7	685	4	BJ062108	BJ062108 BJ062108
37	212.4	25.7	634	1	AL864302	AL864302 AL864302
38	210.2	25.4	715	5	BU337163	BU337163 603514014
39	198.6	24.0	712	5	BX620332	BX620332 BX620332
40	194.4	23.5	666	5	BW215750	BW215750 BW215750
41	191.8	23.2	725	5	BW442024	BW442024 BW442024
42	188.2	22.7	398	6	CB707581	CB707581 AMGNNUC:M
43	186	22.5	325	6	CB298572	CB298572 220014_re
44	180.4	21.8	741	5	BW216493	BW216493 BW216493
45	178.2	21.5	1562	3	AK075884	AK075884 Mus muscu

ALIGNMENTS

RESULT 1
AY401889
LOCUS AY401889 681 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION AY401889
VERSION AY401889.1 GI:39757875
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 681)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Farrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 681)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Farrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..681
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>681
/locus_tag="HCM1041"

Query Match 81.1%; Score 671.4; DB 9; Length 681;
Best Local Similarity 99.1%; Pred. No. 2.7e-180;
Matches 675; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 148 ATGTGCTTGTGTGATGAGTGAACATGCTTGACAGTGCACATCTACTTAACTCGACT 207
Db 1 ATGTGCTTGTGTGATGAGTGAACATGCTTGACAGTGCACATCTACTTAACTCGACT 60

KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

EST.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 1022)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LAM14067 Row: j Column: 05
High quality sequence stop: 681.
Location/Qualifiers
1..1022
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6505924"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_134"
/note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
this is a NIH_MGC Library."

FEATURES
source
1..1022
71.7%; Score 593.6; DB 5; Length 1022;
Best Local Similarity 85.5%; Pred. No. 5.1e-158;
Matches 708; Conservative 0; Mismatches 114; Indels 6; Gaps 4;

ORIGIN
Query Match
Best Local Similarity 85.5%; Pred. No. 5.1e-158;
Matches 708; Conservative 0; Mismatches 114; Indels 6; Gaps 4;
Qy 1 ATGGGCGCCGCACTGGTGGGACCAAGCTGCAGGCTGGTAGCTCGGAGTGGAGCTGTGC 60
Db 120 ATGGGCGCCGCACTGGTGGGACCAAGCTGCAGGCTGGTAGCTCGGAGTGGAGTGTGC 179
Qy 61 GAGGACACTACACATCGTGGCTGCTATCGCCAGTCTTCAACACGATCAGCAATGTC 120
Db 180 GAGGACACTACATCGTGGCTGCTATCGCCAGTCTTCAACACGATCAGCAATGTC 239
Qy 121 TTATTTTTCATTTTACCGCCCATCTGCATGTGCTTGTGTGATGATGATGCAACATGCTTG 180
Db 240 TTGTTTTCATTTTACCTCCCATCTGCATGTGCTTGTTCGCCAGTACGCAACGTGCTTC 299
Qy 181 AACAGTGACATCTAATACTGAATCTTTTGGTGTGAGTGGGAATGGATCGCTCTAC 240
Db 300 AACAGTGACATCTAATACTGAATCTTTTGGTGTGAGTGGGAATGGATCGCTCTAC 359
Qy 241 TTCCATTTTACCTTAGTTTCTTGGGTGAGTGTGATGATGATGATGATGATGATGATGATG 300
Db 360 TTCCATTTTACCTTAGTTTCTTGGGTGAGTGTGATGATGATGATGATGATGATGATGATG 419
Qy 301 CTGATGTGTGCTTGGCCATGTGTTTCCCGAGAGGTATCTACCAAGATCTTTTCGGAAT 360
Db 420 CTGATGTGTGCTTGGCCATGTGTTTCCCGAGAGGTATTTACCAAGATCTTTTCGGAAT 479
Qy 361 GACAGGGTAGTTCAGAGTGGTGGTTCAGTGTCTCTGCTGCGGTTTACGAGTGGCTGSCA 420
Db 480 GACAGGGTAGTTCAGAGTGGTGGTTCAGTGTCTCTGCTGCGGTTTACGAGTGGCTGSCA 539
Qy 421 TTTGTCAAGCTGCGATCAACACATCTCTGTGATGACCCCTGGGAGTTTCTTGGCACTGCA 480
Db 540 TTTATCAAGCCCGCATCAACAATATTTTCCGTGATGATTTCTGGGACTTCCATGCACTGCG 599
Qy 481 CTGCTCATCGCAGAGCTAAAGAGGTGTGACAAATGCGTGTGTTTAAAGCTGGGCTCTTC 540

Db 600 CTGCTTGTTCAGAGCTGAAGAGGTGTGACAAATGTGCTGTGTTTAAAGCTGGGCTCTTC 659
Qy 541 TCGGGCTCTCTGGTGGACCTTGGCCCTGTGCTGGATCAGTACCGAGCTTTCTGGCA 599
Db 660 TCTGGCTCTCTGGTGGACCTTCTGCTGGATCAGCAGCAAGCCTTCTGTGA 719
Qy 600 GCTGCTGTATCTTCAACTTCCCTTACCTGATGATGATGATGATGATGATGATGATGATG 659
Db 720 GCTGCTGTATCTTCAACTTCCCTTACCTGATGATGATGATGATGATGATGATGATGATG 779
Qy 660 TGCTGCTTACCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 719
Db 780 TGCTGCTTACCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 839
Qy 720 GCAAGGCTCTGTATCAAG--TTCTGGCCCAATGAGAAATGAGGCT--TTCAATGCTGTGC 775
Db 840 GCAAGTCCAGTCCATCAGATTCCTGGGCCCAAGGAGGAAATGAGGCTTTTATTTGGGGTCC 899
Qy 776 CTAATGTGCTCTCTCTGTG--CCAAAGAAATCATCAGTCAAGATC 822
Db 900 CTAATGTGCTCTCTCTGTG--CCAAAGAAATCATCAGTCAAGATC 947

RESULT 4
LOCUS AY401891 681 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM1041 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY401891
VERSION AY401891.1 GI:39757877
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 681)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarial, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED 14671302
REFERENCE
AUTHORS
2 (bases 1 to 681)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarial, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
TITLE
JOURNAL
COMMENT
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..681
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>681
/locus_tag="HCM1041"

gene
ORIGIN
Query Match
Best Local Similarity 86.6%; Pred. No. 1.9e-141;
Matches 590; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
Qy 148 ATGTCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
Db 1 ATGTCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
Qy 208 CTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20493374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 1173)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/
FEATURES	Location/Qualifiers
source	1..1173 /organism="Mus musculus"


```
|||||
Db 1 TAGTGGGAATCGGATCTGTCTACTTCCATNGCAACTCTTAGTTTCTGGGTCAGATGCTT 60
QY 277 GATCAACTTGCAGTCCCTTTGGTGTCTGATGTGTCTTTGGCCATGTGTTCCCCAGAGG 336
Db 61 GATGAACCTTGCATCTTTGGTGTCTGATGTGTCTTTGGCCATGTGTTCCCCAGAGG 120
QY 337 TATCTACCAAGATCTTTTCGGAATGACAGGGGTAGGTTCAAGGTGGTGGTTCAGTGTCTTG 396
Db 121 TATTTACCAAGATCTTTTCGGAATGACAGGGGTAGGTTCAAGGTGGTGGTTCAGTGTCTTG 180
QY 397 TCTCGGTTAGCAGTGGTGGTCTTTTCGAGCTGCTGATCAAGCTGCTGATCAAGCTGCTGATG 456
Db 181 TCTCGAGTTACCAATGCTGGCTTTTGTCAAGCTGCTGATCAAGCTGCTGATCAAGCTGCTGATG 240
QY 457 ACCCTGGGAGTTCCTTTCGACATGCACTGCTCATCGCAGAGCTAAAGAGGTGTGACAAATG 516
Db 241 ATTTCTGGGGTTCGTTGACATGCGCTGCTCATTTGAGAGCTGCAAGAGGTGTGACAAATG 300
QY 517 CGTGTGTTTAAAGCTGGGCTCTTCTCGGGCTCTGTTGAGAGCTGCTGATCAAGCTGCTGATG 576
Db 301 CGTGTGTTTAAAGCTGGGCTCTTCTCGGGCTCTGTTGAGAGCTGCTGATCAAGCTGCTGATG 360
QY 577 ATCAGTGACCGAGCTTTCTGAGCTGCTGATCAAGCTGCTGATCAAGCTGCTGATCAAGCTGCTG 636
Db 361 ATCAGTGACCGAGCTTTCTGAGCTGCTGATCAAGCTGCTGATCAAGCTGCTGATCAAGCTGCTG 420
QY 637 ATGTGGACATCTCTCATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 696
Db 421 GTGTGGACATCTCTCATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 480
QY 697 GATGCTGCTCAGAGATCTCTGAGCAAGGCTGCTGATCAAGCTGCTGATCAAGCTGCTGATCAAG 756
Db 481 GATGCTGCTCAGAGATCTCTGAGCAAGGCTGCTGATCAAGCTGCTGATCAAGCTGCTGATCAAG 540
QY 757 TGGGCTTCATGTGTCCTCTATG 781
Db 541 TGGGCTTTTATCGGTGTCCTCTATG 565
```

```
RESULT 9
BX646596 449 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp781B0790_r1.781 (synonym: hlcc4) Homo sapiens cDNA clone
DEFINITION DKFZp781B0790 5', mRNA sequence.
ACCESSION BX646596
VERSION BX646596.1 GI:34480929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 449)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Foboh, G., Han, M., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
JOURNAL EST (Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., et al.)
COMMENT Unpublished (2003)
Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No si sequence available.
This clone (DKFZp781B0790) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source Location/Qualifiers
1..449
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="DKFZp781B0790"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/nore="vector: pSport1_sfi; Site_1: SfiIA; Site_2: SfiIB;
cdna-collection"

ORIGIN
Query Match 54.0%; Score 447.4; DB 5; Length 449;
Best Local Similarity 99.8%; Pred. No. 2.3e-116;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 265 GGTGAGATGCTGTGATGAACCTTTCGAGTCCCTTTGGGTTCTGATGTGCTTTGGCCATGCG 324
Db 1 GGTGAGATGCTGTGATGAACCTTTCGAGTCCCTTTGGGTTCTGATGTGCTTTGGCCATGCG 60
QY 325 TTCCCGAAGGATCTACCAAGATCTTTCGGAATGACAGGGGTAGGTTCAAGGTGGTG 384
Db 61 TTCCCGAAGGATCTACCAAGATCTTTCGGAATGACAGGGGTAGGTTCAAGGTGGTG 120
QY 385 GTCAGTGTCTGTCTGCGGTTTACGACGTGCGCTTGTCAAGCTTGTCAAGCTTGTCAAGCTTGT 444
Db 121 GTCAGTGTCTGTCTGCGGTTTACGACGTGCGCTTGTCAAGCTTGTCAAGCTTGTCAAGCTTGT 180
QY 445 ATCTCTCTGATGACCTCGGAGTTCCTTTCGAGCTGCTGATGCTGATGCTGATGCTGATGCTG 504
Db 181 ATCTCTCTGATGACCTCGGAGTTCCTTTCGAGCTGCTGATGCTGATGCTGATGCTGATGCTG 240
QY 505 TGTGACACATGCGGTGCTTTAAAGCTGGGCTCTTTCGAGCTGCTGATGCTGATGCTGATGCTG 564
Db 241 TGTGACACATGCGGTGCTTTAAAGCTGGGCTCTTTCGAGCTGCTGATGCTGATGCTGATGCTG 300
QY 565 CTGTTCTCTGATGACGTCGAGTTCGAGCTTTCGAGCTGCTGATGCTGATGCTGATGCTGATG 624
Db 301 CTGTTCTCTGATGACGTCGAGTTCGAGCTTTCGAGCTGCTGATGCTGATGCTGATGCTGATG 360
QY 625 TACCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 684
Db 361 TACCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 685 TTTGCTACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 713
Db 421 TTTGCTACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449

RESULT 10
BX660847 633 bp mRNA linear EST 26-OCT-2001
LOCUS BB660847 RIKEN full-length enriched, 0 day neonate kidney Mus
DEFINITION musculus cDNA clone D630008P07 5', mRNA sequence.
ACCESSION BB660847
VERSION BB660847.1 GI:16494626
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 633)
AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Konno, H., Kouda, M., Koyama, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
```


/note=Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus,
semintendous muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

ORIGIN

Query Match 50.3%; Score 416.2; DB 4; Length 498;
Best Local Similarity 91.1%; Pred. No. 2e-107;
Matches 453; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 2 TGGGCGCCCGCACTGGTGGGACAGCTGAGGCTGGTAGCTCGAGGTGGACTGGTGG 61
DB 1 TGAGCGCCCGCACTGGTGGGACAGCTGAGGCTGGTAGCTCGAGGTGGACTGGTGG 60

QY 62 AGGACAACTACACCATGGTGGGCTGCTATCGCGAGTCTCAACACCATGAGCAATGCT 121
DB 61 AGGACAACTACACCATGGTGGGCTGCTATCGCGAGTCTCAACACCATGAGCAATGCT 120

QY 122 TATTTTTCATTTTACCGCCCATCTGCGATGCTGCTTGTGATGAGTATGCAACATGCTGA 181
DB 121 TGTTTTTCATCTTGGCGCCCATCTGCGATGCTGCTTGTGATGAGTATGCAACATGCTGA 180

QY 182 ACAGTGACATCTAATCTGCACTCTTTGGTGTGATGAGTATGCAACATGCTGCTACT 241
DB 181 ACAGTGACATCTAATCTGCACTCTTTGGTGTGATGAGTATGCAACATGCTGCTACT 240

QY 242 TCCATTTTACCTTTAGTTTCTGGTGCAGATGCTTGTGATGAGTATGCAACATGCTGCTTC 301
DB 241 TCCATTTTACCTTTAGTTTCTGGTGCAGATGCTTGTGATGAGTATGCAACATGCTGCTTC 300

QY 302 TGATGTGCTTGGTGCAGATGCTTGTGATGAGTATGCAACATGCTGCTTC 361
DB 301 TGATGTGCTTGGTGCAGATGCTTGTGATGAGTATGCAACATGCTGCTTC 360

QY 362 ACAGGGGTAGTTCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 421
DB 361 ACAGGGGTAGTTCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420

QY 422 TTGTCAAGCTGCTTACACATCTCTCTCATGACC-CTGGGAGTTCTTGTGCACTGCA 480
DB 421 TTGTCAAGCTGCTTACACATCTCTCTCATGACCCTGGGAGTTCTTGTGCACTGCA 480

QY 481 CTGCTCATCGCAGAGCT 497
DB 481 CTACTCATGCTGAGCT 497

RESULT 12
AF370405 1527 bp mRNA linear HTC 31-DEC-2003
LOCUS Homo sapiens P11646 mRNA, complete cds.
DEFINITION AF370405
ACCESSION AF370405.1 GI:33341739
VERSION AF370405.1
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1527)
AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,
Wan, D.F. and Gu, J.R.
Novel human cDNA clones with function of inhibiting cancer cell
growth
TITLE Unpublished
JOURNAL 2 (bases 1 to 1527)
REFERENCE Zhang, P.P., Zhou, X.M., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T.,
AUTHORS Wan, D.F. and Gu, J.R.
Direct Submission
TITLE Submitted (16-APR-2001) National Laboratory for Oncogenes & Related
JOURNAL Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
200032, P. R. China
FEATURES Location/Qualifiers
1. .1527
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
218. .640
/codon_start=1
/product="P11646"
/protein_id="AA015241.1"
/db_xref="GI:33341740"
/translation="MCLFROYATCFNSGIYLIWLLVVGIGSVYPHATLSFLGQMLD
ELAVLWLMCALAMWPPRRYLPKIFRNDGRKFKVSVLSAVTICLAFVPAINNISL
MTLGVPTALLIAELKHERNRRHRKGGQGGGDKV"

ORIGIN

Query Match 50.1%; Score 414.8; DB 3; Length 1527;
Best Local Similarity 82.4%; Pred. No. 7e-107;
Matches 532; Conservative 0; Mismatches 17; Indels 97; Gaps 1;

QY 108 GATCAGCAATGCTTATTTTTCATTTTACCGCCCATCTGCAATGCTGCTTGTGATGAGT 167
DB 178 GATCAGCAATGCTTATTTTTCATTTTACCGCCCATCTGCAATGCTGCTTGTGATGAGT 237

QY 168 TGCACATGCTTGAACAGTGCATCTACTTAATCTGCACTCTTTTGTGATGAGTGGAA 227
DB 238 TGCACATGCTTGAACAGTGCATCTACTTAATCTGCACTCTTTTGTGATGAGTGGAA 297

QY 228 TGGATCGCTTACTTCCATTTTACCGCCCATCTGCAATGCTGCTTGTGATGAGTGG 287
DB 298 TGGATCGCTTACTTCCATTTTACCGCCCATCTGCAATGCTGCTTGTGATGAGTGG 357

QY 288 AGTCTTTGGGTTCTGATGCTGCTTGGCCATGCTGCTTGGCCATGCTTGGCCATGCT 347
DB 358 AGTCTTTGGGTTCTGATGCTGCTTGGCCATGCTGCTTGGCCATGCTTGGCCATGCT 417

QY 348 GATCTTTCCGAATGACAGGGTAGGTTCAAGGTGGTGGTGGTGGTGGTGGTGGTGG 407
DB 418 GATCTTTCCGAATGACAGGGTAGGTTCAAGGTGGTGGTGGTGGTGGTGGTGGTGG 477

QY 408 GACGTGCTTGGCATTTGTGAGCCCTGCAATCAACAACTCTCTGATGACCCCTGGAG 467
DB 478 GACGTGCTTGGCATTTGTGAGCCCTGCAATCAACAACTCTCTGATGACCCCTGGAG 537

QY 468 TCCTTGGCATGCTGCTCATCGCAGAGCTAAG----- 501
DB 538 TCCTTGGCATGCTGCTCATCGCAGAGCTAAGAGGCGATGAGAGAACCGAGAGAG 597

QY 502 ----- 501

DB 598 ACACAGGAAGGTGGCCAGAGAGGTGGAGCAAGGTCTGACGATGAGTGACTCTCTG 657

QY 502 -----AGGTGTGACACATCGCTGCTTTAAGCTGGGCTCTTCTCGGGCTCT 550

DB 658 GACCCCGTGCAGGTGTGACACATCGCTGCTTTAAGCTGGGCTCTTCTCGGGCTCT 717

QY 551 GGTGGACCTGGCCCTGTTCTGCTGATGAGTGCAGCCAGCTTTCTCGAGCTGCTGCT 610

DB 718 GGTGGACCTGGCCCTGTTCTGCTGATGAGTGCAGCCAGCTTTCTCGAGCTGCTGCT 777

QY 611 CCTTCAACTTCCCTACCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656

DB 778 CCTTCAACTTCCCTACCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823

RESULT 13
BU234223 698 bp mRNA linear EST 26-NOV-2002
LOCUS BU234223
DEFINITION BU234223 CSQCHN24 Gallus gallus cDNA clone ChEST75k21 5', mRNA
sequence.
ACCESSION BU234223
VERSION BU234223.1 GI:25478587
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mon Nov 22 17:00:24 2004

```

Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0812 Row: G Column: 07
Seq primer: M13 Reverse
High quality sequence stop: 548
POLYA=No.

FEATURES
    source
    Location/Qualifiers
        1..548
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="niaEST:B0812G07-5"
            /clone="NIA:B0812G07 IMAGE:30468654"
            /dev_stage="Newborn Kidney"
            /lab_host="DHI08"
            /clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
            1)"
            /note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
            Site_2: NotI; Mouse cDNA project by the Laboratory of
            Genetics, National Institute on Aging (NIA), Intramural
            Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
            In brief, double-stranded cDNAs were synthesized with an
            Oligo(dT) primer [Invitrogen:
            5'-TGACTAGTCTAGATCGGAGCGCCGCTTTTCTTTT-3'] from
            26 ug of total RNA, treated with T4 DNA polymerase, and
            purified by ethanol-precipitation. The cDNAs were ligated
            to lona-linker LL-Sal4, purified by phenol/chloroform, and
            separated from free linkers by Centricon 100. Then, the
            cDNAs were amplified by long-range high fidelity PCR using
            Ex Taq polymerase (Takara) with a primer Sal4-S. The
            products were purified by phenol/chloroform and Centricon
            100. The cDNAs were digested with SalI and NotI enzymes
            and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
            vector. The DHI08 E. coli host was transformed with the
            ligation mixture by the standard chemical method. The
            average insert size is about 3.0 kb. The library was
            constructed by Yulan Piao."

Query Match 45.0%; Score 372.6; DB 6; Length 548;
Best Local Similarity 85.7%; Pred. No. 5.8e-95;
Matches 414; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 22 GACCAGCTCGAGCTGGTAGCTCGGAGGTGGAGTGGTGGAGGACAACTACACATCGTG 81
Db 1 GACCAGCTCGGAGCTGGTAGCTCGGAGGTGGAGTGGTGGAGGACAACTACACATCGTG 60
QY 82 CTGCTATCGCCGAGTTCTACACAGATCAGCAATGCTTTATTTTCTATTTTACCGCCC 141
Db 61 CTGCGCATTCGCGAGTTCTACACAGATCAGCAATGCTTTATTTTCTATTTTACCTCCC 120
QY 142 ATCTGCAATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 201
Db 121 ATCTGCAATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 202 TGAATCTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 261
Db 181 TGAATCTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
QY 262 TTGGGTTCAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 321
Db 241 CTGGGTTCAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
QY 322 TGGTTCCTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 381
Db 301 TGGTTCCTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
QY 382 GTGTTCAGTTCCTGCTGCGGTTACGACGTGCTGGCAATTTGTCAAGCCCTGCCATCAAC 441
Db 361 GTGTTCAGTTCCTGCTGCGGTTACGACGTGCTGGCAATTTGTCAAGCCCTGCCATCAAC 420
QY 442 AACATCTCTGTGATGACCCCTGGGAGTTCTTGTGCACTGCACTGCTCATCGAGAGCTAAG 501
Db 421 AACATCTCTGTGATGACCCCTGGGAGTTCTTGTGCACTGCACTGCTCATCGAGAGCTAAG 480

ORIGIN
Query Match 45.0%; Score 372.6; DB 6; Length 548;
Best Local Similarity 85.7%; Pred. No. 5.8e-95;
Matches 414; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 22 GACCAGCTCGAGCTGGTAGCTCGGAGGTGGAGTGGTGGAGGACAACTACACATCGTG 81
Db 1 GACCAGCTCGGAGCTGGTAGCTCGGAGGTGGAGTGGTGGAGGACAACTACACATCGTG 60
QY 82 CTGCTATCGCCGAGTTCTACACAGATCAGCAATGCTTTATTTTCTATTTTACCGCCC 141
Db 61 CTGCGCATTCGCGAGTTCTACACAGATCAGCAATGCTTTATTTTCTATTTTACCTCCC 120
QY 142 ATCTGCAATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 201
Db 121 ATCTGCAATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 202 TGAATCTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 261
Db 181 TGAATCTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
QY 262 TTGGGTTCAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 321
Db 241 CTGGGTTCAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
QY 322 TGGTTCCTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 381
Db 301 TGGTTCCTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
QY 382 GTGTTCAGTTCCTGCTGCGGTTACGACGTGCTGGCAATTTGTCAAGCCCTGCCATCAAC 441
Db 361 GTGTTCAGTTCCTGCTGCGGTTACGACGTGCTGGCAATTTGTCAAGCCCTGCCATCAAC 420
QY 442 AACATCTCTGTGATGACCCCTGGGAGTTCTTGTGCACTGCACTGCTCATCGAGAGCTAAG 501
Db 421 AACATCTCTGTGATGACCCCTGGGAGTTCTTGTGCACTGCACTGCTCATCGAGAGCTAAG 480

RESULT 15
CF169369 548 bp mRNA linear EST 25-JUL-2003
LOCUS B0812G07-5 NIA Mouse Newborn Kidney cDNA Library (Long 1) Mus
DEFINITION Musculus cDNA clone NIA:B0812G07 IMAGE:30468654 5', mRNA sequence.
ACCESSION CF169369
VERSION CF169369.1 GI:33278918
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 548)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
2142098
11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

```

Db 421 AATATTTCCTGATGATTCTGGGACTTCATGCACTGCGCTGCTTTGTCAGAGCTGAAG 480
Qy 502 AGG 504
Db 481 AGG 483

Search completed: November 21, 2004, 06:57:34
Job time : 2289.57 secB